

1998

# On analyzing GMRES from the view of geometry and an application to Haar wavelets

Mei-Wern Cheng  
*San Jose State University*

Follow this and additional works at: [https://scholarworks.sjsu.edu/etd\\_theses](https://scholarworks.sjsu.edu/etd_theses)

---

## Recommended Citation

Cheng, Mei-Wern, "On analyzing GMRES from the view of geometry and an application to Haar wavelets" (1998). *Master's Theses*. 1690.

DOI: <https://doi.org/10.31979/etd.24bd-tbdh>

[https://scholarworks.sjsu.edu/etd\\_theses/1690](https://scholarworks.sjsu.edu/etd_theses/1690)

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact [scholarworks@sjsu.edu](mailto:scholarworks@sjsu.edu).

## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

# UMI

A Bell & Howell Information Company  
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA  
313/761-4700 800/521-0600



## **NOTE TO USERS**

**The original document received by UMI contains pages with slanted print. Pages were microfilmed as received.**

**This reproduction is the best copy available**

**UMI**



**ON ANALYZING GMRES FROM THE VIEW OF GEOMETRY**  
**and an application to Haar wavelets**

**A Thesis**

**Presented to**

**The Faculty of the Department of Mathematics**

**San Jose State University**

**In Partial Fulfillment**

**of the Requirements for the Degree**

**Master of Science**

**By**

**Mei-Wern Cheng**

**August 1998**

**UMI Number: 1391513**

**Copyright 1998 by  
Cheng, Mei-Wern**

**All rights reserved.**

---

**UMI Microform 1391513  
Copyright 1998, by UMI Company. All rights reserved.**

**This microform edition is protected against unauthorized  
copying under Title 17, United States Code.**

---

**UMI**  
300 North Zeeb Road  
Ann Arbor, MI 48103

© 1998

Mei-Wern Cheng


ALL RIGHTS RESERVED



APPROVED FOR THE DEPARTMENT OF MATHEMATICS



Dr. Howard Swann




Dr. Leslie Foster



Dr. Mohammad Saleem

APPROVED FOR THE UNIVERSITY



## **ABSTRACT**

### **ON ANALYZING GMRES FROM THE VIEW OF GEOMETRY and an application to Haar wavelets**

**by Mei-Wern Cheng**

Two recently developed Krylov method for solving linear systems are Arnoldi's method and the Generalized Minimum Residual (GMRES) method. We show here that there is a relationship between breakdowns in the two methods. We explore the GMRES from the view of geometry when Arnoldi's algorithm breaks down. The result in this paper suggests that the choice of the initial guess plays a very important role in deciding how the GMRES method will converge or when the GMRES method will converge.

"Wavelets" or "wavelet transforms" are a tool for decomposing functions in various applications. In this paper, we choose the simplest wavelet, Haar wavelet, to introduce the wavelet transform and apply the GMRES method to find the wavelet transform coefficients. We show a discrete data set can be decomposed by a wavelet transform, then de-noised, and reconstructed, from the beginning to the end, with the assistance of the GMRES method.

## **ACKNOWLEDGMENTS**

**I would like to thank Dr. Mohammad Saleem for the guidance given during the course of writing this thesis and for choosing this interesting topic, which leads me to the new world of applied mathematics.**

**Next, I wish to express my appreciation to Dr. Howard Swann, my life time professor (if he has not wanted to give me up yet). I have taken four classes with him because .....(because he is Dr. Swann). If I count the previous time he has devoted to read, discuss and correct my thesis I am taking my fifth class with him. I know that it is not an enjoyable task reading my thesis, but Dr. Swann with his patience, tried to make this thesis more readable. His signature of approval in my thesis is priceless to me.**

**I would also like to thank Dr. Leslie Foster for being on the Comprehensive Oral Examination Committee. His comments and suggestions have been valuable.**

**A special thanks is for Dr. Hedley Morris who spent many hours in introducing me to the field of wavelets. I appreciate his precious time and wisdom.**

**And a lot of thanks to Gary Hui, one of Dr. Swann's thesis students, for his kindness in teaching me how to use Microsoft word equation editor.**

**Being a student and a mother with two children at the same time, I try to be more of a mother than a student - I want straight As in my sons' report cards more than in my report cards. Indeed, I have to thank my husband's support and my two sons' cooperation so that we all have straight As (except my English 100W), and I need not to feel guilty of being a "student mother".**

**It is to my teachers, friends, and my family that I dedicate this manuscript.**

## CONTENT

List of figures.....	viii
List of tables.....	ix
Introduction.....	1.
Chapter 1. Arnoldi Method and GMRES method.....	3.
Section 1.1 Projection concept.....	3.
Section 1.2 Krylov subspace method.....	5.
Section 1.3 Overview of two methods.....	6.
Section 1.4 Arnoldi's Method.....	8.
Section 1.5 GMRES Method.....	14.
Section 1.6 Some conclusions of the case when $\mathbf{H}_m$ becomes singular.....	22.
Section 1.7 Cases that make $\mathbf{H}_m$ become singular.....	24.
Section 1.8 proofs that $\mathbf{x}_{m-1}^g = \mathbf{x}_m^g$ in some cases.....	26.
8.1 $\mathbf{A}\mathbf{v}_m = 0$ .....	27.
8.2 $\mathbf{A}\mathbf{v}_m \perp \mathbf{v}_i$ for $i = 1, 2, 3, \dots, m$ .....	32.
8.3 The $m^{\text{th}}$ column of $\mathbf{H}_m$ is a linear combination of the previous columns.....	40.
Section 1.9 Restarting GMRES.....	44.
Section 1.10 Conclusions.....	48.
Chapter 2. Wavelets.....	53.
Section 2.1 An introduction to wavelets.....	53.

<b>Section 2.2 Haar basis.....</b>	<b>56.</b>
<b>Section 2.3 Wavelet transforms for discrete data.....</b>	<b>61.</b>
<b>Section 2.4... Conclusion and further studies.....</b>	<b>71.</b>
<b>Appendix A.....</b>	<b>72.</b>
<b>Appendix B.....</b>	<b>73.</b>
<b>Bibliography.....</b>	<b>74.</b>

**LIST of FIGURES:**

**Figure 1.1.1 Interpretation of the orthogonality condition..... 4.**

**Figure 1.2.1 Interpretation of the orthogonality condition.....5.**

**Figure 1.3.1 Arnoldi’s method..... 6.**

**Figure 1.3.2 GMRES method..... 6.**

**Figure 1.3.3 A combination of Arnoldi’s and GMRES method..... 7.**

**Figure 1.4.1 Arnoldi’s method for first residual.....9.**

**Figure 1.4.2 Arnoldi’s method for second residual.....9.**

**Figure 1.4.3 Arnoldi’s method for the  $m^{\text{th}}$  residual..... 11.**

**Figure 1.4.4  $h_{m+1,m} = \| \mathbf{u}_{m+1} \|$ ..... 11.**

**Figure 1.5.1 GMRES method for first residual..... 15.**

**Figure 1.5.2 GMRES method for second residual..... 15.**

**Figure 1.5.3 GMRES method for the  $m^{\text{th}}$  residual..... 16.**

**Figure 1.5.4  $h_{m+1,m} = \| \mathbf{u}_{m+1} \|$ ..... 16.**

**Figure 1.8.1 Pythagorean Theorem in case (ii)..... 33.**

**Figure 2.2.1 Wavelets.....60.**

**Figure 2.2.2 Discrete data..... 60.**

**Figure 2.2.3 Averages and sums of the discrete data..... 60.**

**Figure 1. for appendix B.....73.**

**LIST of TABLES:**

<b>Table 1: Three samples of matrix A.....</b>	<b>64.</b>
<b>Table 2: A program to construct matrix A.....</b>	<b>65.</b>
<b>Table 3: Original data vector and picture.....</b>	<b>67.</b>
<b>Table 4: The wavelet transform coefficients before de-noise.....</b>	<b>68.</b>
<b>Table 5: The wavelet transform coefficients after de-noise.....</b>	<b>69.</b>
<b>Table 6: The new data and picture after “de-noise” .....</b>	<b>70.</b>

## **INTRODUCTION**

Two recently developed Krylov methods for solving linear systems  $\mathbf{Ax} = \mathbf{b}$  are Arnoldi's method [4] and the Generalized Minimum Residual (GMRES) method [1]. In the present paper we introduce both Arnoldi's method and the GMRES method. The GMRES method has been considered superior to Arnoldi's method [2]. However, the GMRES method is based on the Arnoldi process. The advantage of these two methods is that we can solve  $n \times n$  systems of linear equations by using smaller systems.

Based on the Arnoldi process, we discuss how the choice of the initial guess  $\mathbf{x}_0$  plays a very important role in the GMRES method. We adopt the result from [2]; but in this paper we prove and explain the result of GMRES according to its geometry. From the view of geometry, it is obvious that for some initial guesses the GMRES algorithm with periodic restarts, which we describe in section 1.9, will not converge. Examples and programs are provided in this paper.

If the initial residual,  $\mathbf{b} - \mathbf{Ax}_0 = \mathbf{r}_0$ , happens to be an eigen-vector of the matrix  $\mathbf{A}^k$ , then the GMRES method will converge at exactly the  $k^{\text{th}}$  iteration. In addition to studying the residuals, we discuss how the norm of the matrix  $\mathbf{A}$  plays a role in the convergence of the GMRES method. A good application of the GMRES method is solving the coefficients of Haar Wavelets transforms [3] [6] [7] [8].

The wavelet transform is a tool for carving up functions or data into components of different frequency, which allows us to study each component separately.



Wavelet analysis may be thought of as a generalization of analysis by the Hilbert space method, wherein one forms an orthogonal basis on the space of interest. Equations in that space may then be solved in terms of an orthogonal basis [10].

Wavelet analysis is similar to Fourier analysis in the sense that it breaks a signal down into its constituent parts for analysis. Whereas the finite Fourier sine transform breaks the signals into a series of sine waves of different frequencies, the wavelets transform breaks the signal into a series of “wavelets”, which are scaled and shifted versions of the “mother wavelet”. In comparison to the series of sine waves which is smooth and of infinite length, the wavelet is irregular in shape and compactly supported.[9]

A linear data set can be viewed as a function, then approximated by a series of wavelets. This gives the compression technique [6] [7] [11] that can be used for de-noising the signal. We de-noise using a simple truncation of all wavelet coefficients larger than some threshold, and by deleting (set to zero) all smaller wavelet coefficients [11].

## **CHAPTER 1. Arnoldi's Method and GMRES Method**

### **Section 1.1 Projection concept**

This section is based on **Iterative Methods For Sparse Linear Systems**, by Yousef Saad. [13]

Most of the existing practical iterative techniques for solving large linear systems of equations utilize a projection process in one way or another. A projection process represents a canonical way for extracting an approximation to the solution of a linear system from a subspace.

Consider the linear system  $\mathbf{Ax} = \mathbf{b}$ , where  $\mathbf{A}$  is an  $n \times n$  real matrix. In this paper, the same symbol  $\mathbf{A}$  is often used to denote the matrix and the linear mapping in  $\mathfrak{R}^n$  that it represents. The idea of projection techniques is to extract an approximate solution to the above problem from a subspace of  $\mathfrak{R}^n$ . If  $\mathbf{K}$  is this subspace of candidate approximants, or search subspace, and if  $m$  is its dimension, then, in general,  $m$  constraints must be imposed to be able to extract such an approximation. A typical way of describing these constraints is to impose  $n-m$  (independent) orthogonality conditions. Specifically, the residual vector defined by approximation  $\mathbf{x}$ ,  $\mathbf{b} - \mathbf{Ax}$ , is constrained to be orthogonal to  $n-m$  linearly independent vectors. This defines a subspace  $\mathbf{L}$  of dimension  $m$ . This simple framework is common to many different mathematical methods and is known as the Petrov-Galerkin conditions.

A projection technique onto the subspace  $\mathbf{K}$  and orthogonal to  $\mathbf{L}$  is a process which finds an approximate solution  $\mathbf{x}'$  to  $\mathbf{Ax} = \mathbf{b}$  by imposing the conditions that  $\mathbf{x}'$  belong to  $\mathbf{K}$  and that the new residual vector be orthogonal to  $\mathbf{L}$ : find  $\mathbf{x}' \in \mathbf{K}$ , such that  $\mathbf{b} - \mathbf{Ax}' \perp \mathbf{L}$ . If we keep repeating this process until we obtain  $n$  linearly independent residuals, which is not always done, then the next residual will be zero, and we will have an exact solution. Note that if  $\mathbf{x}'$  is written in the form  $\mathbf{x}' = \mathbf{x}_0 + \mathbf{z}$ , and the initial residual vector  $\mathbf{r}_0$  is defined as  $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$ , then  $\mathbf{b} - \mathbf{A}(\mathbf{x}_0 + \mathbf{z}) = \mathbf{b} - \mathbf{Ax}' \perp \mathbf{L}$  or  $\mathbf{r}_0 - \mathbf{Az} \perp \mathbf{L}$ .

In other words, the approximate solution can be defined as

$$\begin{aligned} \mathbf{x}' &= \mathbf{x}_0 + \mathbf{z}, \quad \mathbf{z} \in \mathbf{K}, \\ (\mathbf{r}_0 - \mathbf{Az}, \mathbf{w}) &= 0, \quad \forall \mathbf{w} \in \mathbf{L}. \end{aligned} \quad (1.1.1)$$

The orthogonality condition (1.1.1) imposed on the new residual  $\mathbf{r} = \mathbf{r}_0 - \mathbf{Az}$  is illustrated in figure 1.1.1.

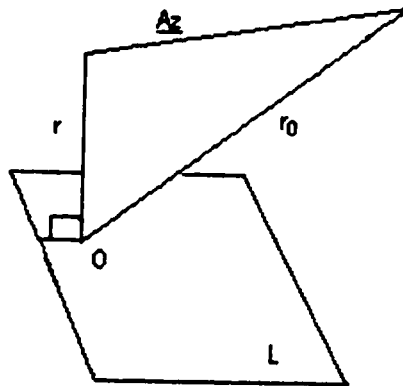


figure 1.1.1 interpretation of the orthogonality condition.

This is a basic projection step, in its most general form. Most standard techniques use a succession of such projections.

**Section 1.2 Krylov subspace methods** (This section is based on *Iterative Methods for sparse linear systems*, by Yousef Saad. [13])

A general projection method for solving the linear system  $\mathbf{Ax} = \mathbf{b}$  is a method which seeks an approximate solution  $\mathbf{x}_m$  from an affine subspace  $\mathbf{x}_0 + \mathbf{K}_m$  of dimension  $m$  by imposing the Petrov-Galerkin condition :  $\mathbf{b} - \mathbf{Ax}_m \perp \mathbf{L}_m$ , where  $\mathbf{L}_m$  is another subspace of dimension  $m$ . Here  $\mathbf{x}_0$  represents an arbitrary initial guess to the solution. A Krylov subspace method is a method for which the subspace  $\mathbf{K}_m$  is the Krylov subspace  $\mathbf{K}_m(\mathbf{A}, \mathbf{r}_0) = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-1}\mathbf{r}_0\}$  where  $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$ . When there is no ambiguity,  $\mathbf{K}_m(\mathbf{A}, \mathbf{r}_0)$  will be denoted by  $\mathbf{K}_m$  (see figure 1.2.1).

Although all the techniques provide the same type of approximations, the choice of  $\mathbf{L}_m$ , i.e., the constraints used to build these approximations, will have an important effect on the iterative technique. We discuss two choices for  $\mathbf{L}_m$ . The first is simply  $\mathbf{L}_m = \mathbf{K}_m$  (for example: Arnoldi's method) and the second is the minimum-residual variation  $\mathbf{L}_m = \mathbf{AK}_m$  for the GMRES method.

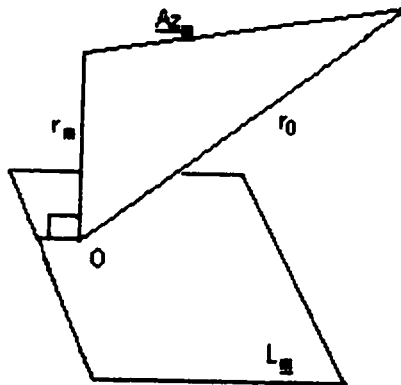


figure 1.2.1. interpretation of the orthogonality condition

### Section 1.3 Overview of Arnoldi's Method and the GMRES Method [2]

If  $\mathbf{x}_0$  is an initial guess for the true solution of  $\mathbf{Ax} = \mathbf{b}$ , then letting  $\mathbf{x}' = \mathbf{x}_0 + \mathbf{z}$ , we have the equivalent system  $\mathbf{r}_0 = \mathbf{Az}$ , where  $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$  is initial residual. Let  $\mathbf{K}_m$  be the Krylov subspace :  $\mathbf{K}_m = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-1}\mathbf{r}_0\}$ . Arnoldi's Method and GMRES Method both find an approximate solution  $\mathbf{x}_m = \mathbf{x}_0 + \mathbf{z}_m$  with  $\mathbf{z}_m \in \mathbf{K}_m$  as shown below:

#### 3.1 Arnoldi's method:

Find  $\mathbf{x}_m$  such that  $(\mathbf{b} - \mathbf{Ax}_m) \perp \mathbf{K}_m$  (equivalently  $(\mathbf{r}_0 - \mathbf{Az}_m) \perp \mathbf{K}_m$ ) (figure 1.3.1 and figure 1.3.3)

Note that :  $\mathbf{r}_m = (\mathbf{b} - \mathbf{Ax}_m) = ((\mathbf{r}_0 - \mathbf{Ax}_0) - \mathbf{A}(\mathbf{x}_0 + \mathbf{z}_m)) = (\mathbf{r}_0 - \mathbf{Az}_m)$

#### 3.2 GMRES method:

Find  $\mathbf{x}_m$  such that  $\|\mathbf{b} - \mathbf{Ax}_m\| = \min \|\mathbf{b} - \mathbf{Ax}\| = \min \|\mathbf{r}_0 - \mathbf{Az}\|$  where  $\mathbf{x} \in \mathbf{x}_0 + \mathbf{K}_m$  and  $\mathbf{z} \in \mathbf{K}_m$ . This is equivalent to requiring  $\mathbf{r}_m = (\mathbf{b} - \mathbf{Ax}_m) = (\mathbf{r}_0 - \mathbf{Az}_m)$  be orthogonal to  $\mathbf{AK}_m$  (figure 1.3.2 and 1.3.3)

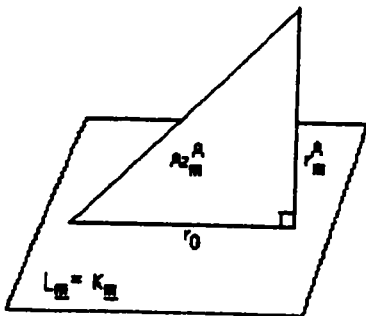


figure 1.3.1 Arnoldi's method

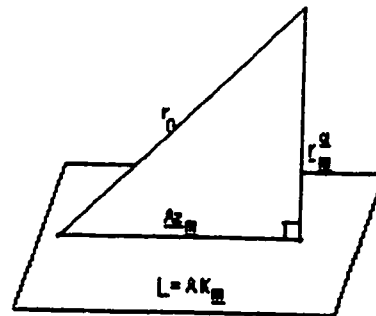
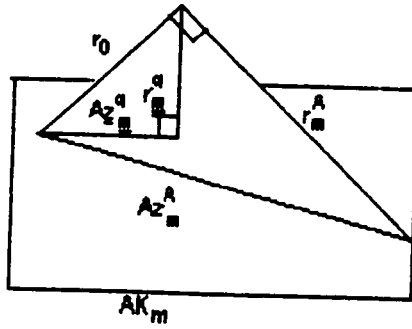


figure 1.3.2 the GMRES method



**figure 1.3.3 A combination of Arnoldi's and GMRES method**

## Section 1.4 Arnoldi's Method

Let  $\mathbf{A}$  be a  $n \times n$  real matrix which is nonsingular, and  $\mathbf{b}$  be a vector  $\in \mathfrak{R}^n$ .

Suppose that  $\mathbf{x}$  is the exact solution for  $\mathbf{Ax} = \mathbf{b}$ . We use the following steps to describe the general concept of Arnoldi's method. The details of computation follow.

4.1. Guess the solution:

Let  $\mathbf{x}_0$  be the initial guess of  $\mathbf{x}$ , then  $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$  will be the initial residual.

Let  $\beta = \|\mathbf{r}_0\|$ , and  $\mathbf{v}_1 = \mathbf{r}_0 / \beta$ .

4.2. Construct the first iteration  $\mathbf{x}_1^A$  and the residual  $\mathbf{r}_1^A$  as follows:

Let  $\mathbf{x}_1^A = \mathbf{x}_0 + \mathbf{z}_1^A$ , where  $\mathbf{z}_1^A \in \mathbf{K}_1 = \text{span}\{\mathbf{r}_0\} = \text{span}\{\mathbf{v}_1\}$ ,

(i.e. there exists scalar  $s$  such that  $\mathbf{z}_1^A = s\mathbf{r}_0$ )

such that  $\mathbf{r}_1^A = \mathbf{b} - \mathbf{Ax}_1^A = \mathbf{r}_0 - \mathbf{Az}_1^A$  is orthogonal to  $\mathbf{r}_0$ .

In other words,  $\mathbf{r}_1^A$  is orthogonal to  $\mathbf{K}_1$  (see figure 1.4.1.)

And  $\mathbf{r}_1^A \in \mathbf{K}_2 = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0\} = \text{span}\{\mathbf{r}_0, \mathbf{r}_1^A\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ , where

$\mathbf{v}_2 = \pm \mathbf{r}_1^A / \|\mathbf{r}_1^A\|$ . Note that  $\mathbf{v}_2$  is orthogonal to  $\mathbf{v}_1$ .

4.3. Construct the second iteration  $\mathbf{x}_2^A$  and the residual  $\mathbf{r}_2^A$  as follows:

Let  $\mathbf{x}_2^A = \mathbf{x}_0 + \mathbf{z}_2^A$ , where  $\mathbf{z}_2^A \in \mathbf{K}_2 = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0\}$ ,

So, there exists scalars  $s$  and  $t$  such that  $\mathbf{z}_2^A = s\mathbf{r}_0 + t\mathbf{Ar}_0$ . Require that

$\mathbf{r}_2^A = \mathbf{b} - \mathbf{Ax}_2^A = \mathbf{r}_0 - \mathbf{Az}_2^A$  be orthogonal to  $\mathbf{r}_0$  and  $\mathbf{r}_1^A$ .

In other words,  $\mathbf{r}_2^A$  is orthogonal to  $\mathbf{K}_2$ . (see figure 1.4.2.)

Let  $\mathbf{v}_3 = \pm \mathbf{r}_2^A / \|\mathbf{r}_2^A\|$ . So  $\mathbf{v}_3$  is orthogonal to  $\mathbf{v}_1$  and  $\mathbf{v}_2$ .

Then  $\mathbf{r}_2^A \in \mathbf{K}_3 = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0\} = \text{span}\{\mathbf{r}_0, \mathbf{r}_1^A, \mathbf{r}_2^A\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ .

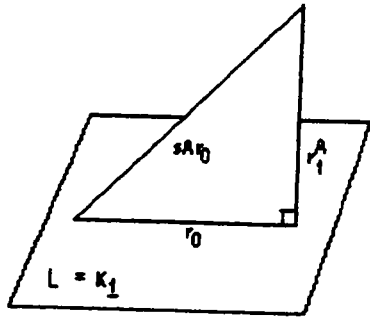


figure 1.4.1

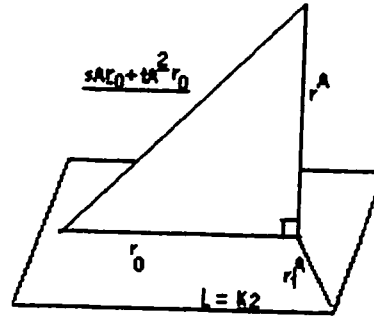


figure 1.4.2

4.4 Proceed inductively at the  $m^{\text{th}}$  iteration:

Construct  $\mathbf{x}_m^A$  and the corresponding residuals  $\mathbf{r}_m^A$ , where  $m = 3, 4, 5, \dots, n$ , by discovering

$\mathbf{x}_m^A = \mathbf{x}_0 + \mathbf{z}_m^A$ , where  $\mathbf{z}_m^A \in \mathbf{K}_m = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-1}\mathbf{r}_0\}$ ,

such that  $\mathbf{r}_m^A = \mathbf{b} - \mathbf{A}\mathbf{x}_m^A = \mathbf{r}_0 - \mathbf{A}\mathbf{z}_m^A$  is orthogonal to  $\mathbf{r}_0, \mathbf{r}_1^A, \mathbf{r}_2^A, \dots$ , and  $\mathbf{r}_{m-1}^A$ .

In other words,  $\mathbf{r}_m^A$  is orthogonal to  $\mathbf{K}_m$ . (figure 1.4.3.)

And  $\mathbf{r}_m^A \in \mathbf{K}_{m+1} = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^m\mathbf{r}_0\}$ .

Let  $\mathbf{v}_{m+1} = \pm \mathbf{r}_m^A / \|\mathbf{r}_m^A\|$ . Then  $\mathbf{v}_{m+1}$  is orthogonal to  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots$ , and  $\mathbf{v}_m$  and  $\text{span}\{\mathbf{r}_0, \mathbf{r}_1^A, \dots, \mathbf{r}_{m-1}^A\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ .



The computational method is to select some  $m$  and use the Gram-Schmidt Process to set  $\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-1}\mathbf{r}_0\}$  to generate orthonormal vectors  $\mathbf{v}_i$  which will be in the direction of  $\mathbf{r}_{i-1}^A$ . The procedure is as follows:

Let  $\mathbf{x}_0$  be the initial guess of  $\mathbf{x}$ . Then  $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$  will be the initial residual.

Let  $\beta = \|\mathbf{r}_0\|$ , and  $\mathbf{v}_1 = \mathbf{r}_0 / \beta$ . Proceed inductively.

Suppose we have  $\mathbf{K}_{m-1} = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-2}\mathbf{r}_0\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_{m-1}\}$ , where  $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_{m-1}\}$  is an orthonormal basis for  $\mathbf{K}_{m-1}$ .

Let  $h_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$  where  $(\ast, \ast)$  denotes inner product. This is defined for  $i \leq j+1$ .

$$\text{Define } \mathbf{u}_m = \mathbf{A}\mathbf{v}_{m-1} - \sum_{i=1}^{m-1} h_{i,m-1} \mathbf{v}_i. \quad (1)$$

$$\text{Then, for } k \leq m-1, (\mathbf{u}_m, \mathbf{v}_k) = (\mathbf{A}\mathbf{v}_{m-1}, \mathbf{v}_k) - \sum_{i=1}^{m-1} h_{i,m-1} (\mathbf{v}_i, \mathbf{v}_k) = (\mathbf{A}\mathbf{v}_{m-1}, \mathbf{v}_k) - h_{k,m-1} = 0.$$

We normalize by letting  $\mathbf{v}_m = \mathbf{u}_m / \|\mathbf{u}_m\|$ . If  $\mathbf{K}_m = \{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-1}\mathbf{r}_0\}$  is an independent set, then  $\mathbf{G}_m = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m\}$  is an orthonormal basis for  $\mathbf{K}_m$ .

Therefore,  $(\mathbf{v}_m, \mathbf{v}_k) = (\mathbf{u}_m / \|\mathbf{u}_m\|, \mathbf{v}_k) = (\mathbf{u}_m, \mathbf{v}_k) / \|\mathbf{u}_m\| = 0 / \|\mathbf{u}_m\| = 0$ .

Note that since  $\mathbf{r}_{m-1}^A = \mathbf{b} - \mathbf{A}\mathbf{x}_{m-1}^A = \mathbf{r}_0 - \mathbf{A}\mathbf{z}_{m-1}^A$  where

$$\mathbf{z}_{m-1}^A \in \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-2}\mathbf{r}_0\},$$

$$\mathbf{r}_{m-1}^A \in \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-1}\mathbf{r}_0\} = \mathbf{K}_m. \text{ Now } \mathbf{r}_{m-1}^A = \mathbf{b} - \mathbf{A}\mathbf{x}_{m-1}^A = \mathbf{r}_0 - \mathbf{A}\mathbf{z}_{m-1}^A$$

was constructed to be orthogonal to  $\mathbf{r}_0, \mathbf{r}_1^A, \dots$ , and  $\mathbf{r}_{m-2}^A \in \mathbf{K}_{m-1}$ , so  $\mathbf{r}_{m-1}^A$  is orthogonal to  $\mathbf{K}_{m-1}$ . Also  $\mathbf{v}_m$  has the same property, hence,  $\pm \mathbf{r}_{m-1}^A / \|\mathbf{r}_{m-1}^A\| = \mathbf{u}_m / \|\mathbf{u}_m\| = \mathbf{v}_m$ . So,

$\mathbf{r}_{m-1}^A$  is orthogonal to  $\mathbf{v}_1, \mathbf{v}_2, \dots$ , and  $\mathbf{v}_{m-1}$ , and  $\mathbf{v}_m$  in the direction of  $\mathbf{r}_{m-1}^A$ .

We show that  $h_{m+1,m} = \|\mathbf{u}_{m+1}\|$ .

*Proof:* Since, by equation (1),  $\mathbf{u}_{m+1} = \mathbf{A}\mathbf{v}_m - \sum_{i=1}^m h_{i,m}\mathbf{v}_i$  (figure 1.4.4.),

$$\begin{aligned} h_{m+1,m} &\equiv (\mathbf{A}\mathbf{v}_m, \mathbf{v}_{m+1}) = (\mathbf{A}\mathbf{v}_m, \mathbf{u}_{m+1}/\|\mathbf{u}_{m+1}\|) \\ &= (\mathbf{A}\mathbf{v}_m, \mathbf{u}_{m+1}) / \|\mathbf{u}_{m+1}\| = (\mathbf{u}_{m+1} + \sum_{i=1}^m h_{i,m}\mathbf{v}_i, \mathbf{u}_{m+1}) / \|\mathbf{u}_{m+1}\| = \|\mathbf{u}_{m+1}\|^2 / \|\mathbf{u}_{m+1}\| \\ &= \|\mathbf{u}_{m+1}\|. \end{aligned}$$

Hence,  $\mathbf{v}_{m+1} = \mathbf{u}_{m+1} / \|\mathbf{u}_{m+1}\| = \mathbf{u}_{m+1} / h_{m+1,m}$ . (2)

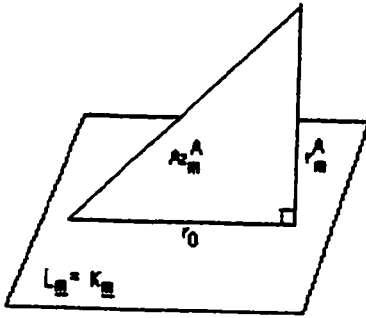


figure 1.4.3 Arnoldi's method

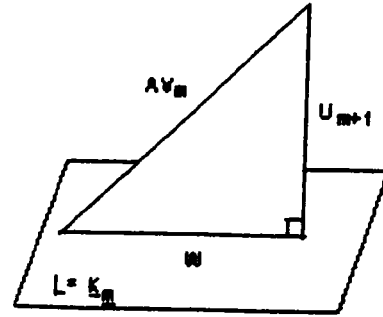


figure 1.4.4 where  $w = \sum_{i=1}^m h_{i,m}\mathbf{v}_i$ .

Define  $\mathbf{H}_m$  to be the  $m \times m$  (Hessenberg) matrix whose entries are the coefficients  $h_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$ ,  $1 \leq i \leq \min\{j+1, m\}$ ,  $1 \leq j$ , and zero elsewhere.

This Hessenberg matrix,  $\mathbf{H}_m$ , is shown in appendix A.

Let  $\mathbf{V}_m = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m]$ , where  $\mathbf{v}_i$  are column vectors representing  $\mathbf{V}_i$ .

Then it is readily seen that  $\mathbf{H}_m = \mathbf{V}_m^T \mathbf{A}\mathbf{V}_m$ . (3)

We show that  $\mathbf{x}_m^A = \mathbf{x}_0 + \beta \mathbf{V}_m \mathbf{H}_m^{-1} \mathbf{e}_m$ , where  $\mathbf{e}_m$  is the column vector  $(1, 0, \dots, 0)^T$ .

*Proof:* For any  $\mathbf{z}_m^A \in \text{span } \mathbf{G}_m = \mathbf{K}_m$ ,  $\mathbf{z}_m^A = \mathbf{V}_m \mathbf{y}_m^A$  for some column vector  $\mathbf{y}_m^A$  with suitable entries.

Since, for  $k = 1 \dots m$ ,  $(\mathbf{v}_k, \mathbf{r}_m^A) = 0$ ,  $\mathbf{V}_m^T \mathbf{r}_m^A = \mathbf{0}$ .

Hence,  $\mathbf{V}_m^T (\mathbf{b} - \mathbf{A} \mathbf{x}_m^A) = \mathbf{0}$ .

Since  $\mathbf{x}_m^A = \mathbf{x}_0 + \mathbf{z}_m^A$ ,  $\mathbf{V}_m^T (\mathbf{b} - \mathbf{A} \mathbf{x}_0 - \mathbf{A} \mathbf{z}_m^A) = \mathbf{0}$ .

So,  $\mathbf{V}_m^T (\mathbf{r}_0 - \mathbf{A} \mathbf{z}_m^A) = \mathbf{0}$  and  $\mathbf{V}_m^T (\mathbf{r}_0 - \mathbf{A} \mathbf{V}_m \mathbf{y}_m^A) = \mathbf{0}$

Therefore,  $\mathbf{V}_m^T \mathbf{r}_0 = \mathbf{V}_m^T \mathbf{A} \mathbf{V}_m \mathbf{y}_m^A$ .

Since, by (3),  $\mathbf{H}_m = \mathbf{V}_m^T \mathbf{A} \mathbf{V}_m$ ,  $\mathbf{V}_m^T \mathbf{r}_0 = \mathbf{H}_m \mathbf{y}_m^A$ .

Therefore, if  $\mathbf{H}_m$  is nonsingular,  $\mathbf{y}_m^A = \mathbf{H}_m^{-1} \mathbf{V}_m^T \mathbf{r}_0$

$$\begin{aligned}
 &= \mathbf{H}_m^{-1} \times [(\mathbf{v}_1, \mathbf{r}_0), (\mathbf{v}_2, \mathbf{r}_0), \dots, (\mathbf{v}_m, \mathbf{r}_0)]_{1 \times m}^T \\
 &= \mathbf{H}_m^{-1} \times [(\mathbf{v}_1, \mathbf{r}_0), 0, \dots, 0]_{1 \times m}^T \\
 &= \mathbf{H}_m^{-1} \times [(\mathbf{v}_1, \beta \mathbf{v}_1), 0, 0, \dots, 0]_{1 \times m}^T \\
 &= \mathbf{H}_m^{-1} \times (\mathbf{v}_1, \beta \mathbf{v}_1) [1, 0, 0, \dots, 0]_{1 \times m}^T \\
 &= \mathbf{H}_m^{-1} \times \beta [1, 0, 0, \dots, 0]_{1 \times m}^T \\
 &= \mathbf{H}_m^{-1} (\beta \mathbf{e}_m).
 \end{aligned}$$

Hence,  $\mathbf{z}_m^A = \mathbf{V}_m \mathbf{y}_m^A = \mathbf{V}_m \mathbf{H}_m^{-1} (\beta \mathbf{e}_m) = \beta \mathbf{V}_m \mathbf{H}_m^{-1} \mathbf{e}_m$ .

Therefore,  $\mathbf{x}_m^A = \mathbf{x}_0 + \beta \mathbf{V}_m \mathbf{H}_m^{-1} \mathbf{e}_m$ , where  $\beta = \|\mathbf{r}_0\|$ . (4)

To summarize, Arnoldi's method is as follows: [2]

First step, start the Gram-Schmidt process:

For an initial guess  $\mathbf{x}_0$ , form  $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$ . Compute  $\beta = \|\mathbf{r}_0\|$ , and let  $\mathbf{v}_1 = \mathbf{r}_0 / \beta$ .

For  $j = 1, 2, \dots, m$  do:

(a) Form  $\mathbf{A}\mathbf{v}_j$  and orthogonalize it against the previous  $\mathbf{v}_1, \dots, \mathbf{v}_j$  via

$$h_{ij} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i), \quad i = 1, 2, \dots, j,$$

$$\mathbf{u}_{j+1} = \mathbf{A}\mathbf{v}_j - \sum_{i=1}^j h_{ij}\mathbf{v}_i.$$

Then  $h_{j+1,j} = \|\mathbf{u}_{j+1}\|$ , and

$$\mathbf{v}_{j+1} = \mathbf{u}_{j+1} / h_{j+1,j}.$$

(b) Compute the residual norm  $\rho_j = \|\mathbf{b} - \mathbf{A}\mathbf{x}_j\|$  of the solution  $\mathbf{x}_j$  that would be obtained if we stopped at this step.

(c) If  $\rho_j \leq \epsilon$  set  $m = j$  and go to the second step.

Second step, form the approximate solution:

Define  $\mathbf{H}_m$  to be the  $m \times m$  (Hessenberg) matrix whose nonzero entries are the coefficients  $h_{ij}$ ,  $1 \leq i \leq \min\{j+1, m\}$ ,  $1 \leq j \leq m$  and define  $\mathbf{V}_m = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m]$ .

Compute  $\mathbf{x}_m^A = \mathbf{x}_0 + \mathbf{z}_m^A$ , where  $\mathbf{z}_m^A = \beta \mathbf{V}_m \mathbf{H}_m^{-1} \mathbf{e}_m$ .

## Section 1.5 GMRES Method

Let  $\mathbf{A}$  be an  $n \times n$  real matrix which is nonsingular, and  $\mathbf{b}$  be a vector  $\in \mathfrak{R}^n$ .

Suppose that  $\mathbf{x}$  is the exact solution for  $\mathbf{Ax} = \mathbf{b}$ .

We use the following steps to describe the general concept of the GMRES method.

5.1. Guess the solution:

Let  $\mathbf{x}_0$  be the initial guess of  $\mathbf{x}$ , then  $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$  will be the initial residual.

Let  $\beta = \|\mathbf{r}_0\|$ , and  $\mathbf{v}_1 = \mathbf{r}_0 / \beta$ .

5.2. Construct the first iteration  $\mathbf{x}_1^g$  and the residual  $\mathbf{r}_1^g$  as follows:

Let  $\mathbf{x}_1^g = \mathbf{x}_0 + \mathbf{z}_1^g$ , where  $\mathbf{z}_1^g \in \mathbf{K}_1 = \text{span}\{\mathbf{r}_0\}$ ,

(So, there exists scalar  $s$  such that  $\mathbf{z}_1^g = s\mathbf{r}_0$ )

such that  $\mathbf{r}_1^g = \mathbf{b} - \mathbf{Ax}_1^g = \mathbf{r}_0 - \mathbf{Az}_1^g$  is orthogonal to  $\mathbf{AK}_1$ .

In other words,  $\mathbf{r}_1^g$  is orthogonal to  $\mathbf{AK}_1 = \text{span}\{\mathbf{Ar}_0\}$ . (see figure 1.5.1.)

And  $\mathbf{r}_1^g \in \mathbf{K}_2 = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0\}$ .

Note that, in the Arnoldi method, the residual  $\mathbf{r}_1^A$  was orthogonal to  $\mathbf{K}_1 = \text{span}\{\mathbf{r}_0\}$ .

5.3. The second iteration  $\mathbf{x}_2^g$  and the residual  $\mathbf{r}_2^g$ :

Let  $\mathbf{x}_2^g = \mathbf{x}_0 + \mathbf{z}_2^g$ , where  $\mathbf{z}_2^g \in \mathbf{K}_2 = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0\}$ ,

(So, there exists scalar  $s$  and  $t$  such that  $\mathbf{z}_2^g = s\mathbf{r}_0 + t\mathbf{Ar}_0$ )

such that  $\mathbf{r}_2^g = \mathbf{b} - \mathbf{Ax}_2^g = \mathbf{r}_0 - \mathbf{Az}_2^g$  is orthogonal to  $\mathbf{AK}_2$ . (see figure 1.5.2.)

And  $\mathbf{r}_2^g \in \mathbf{K}_3 = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0, \mathbf{A}^2\mathbf{r}_0\}$ .

Note that, in the Arnoldi method, residual  $\mathbf{r}_2^A$  was orthogonal to  $\mathbf{K}_2 = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0\}$ .

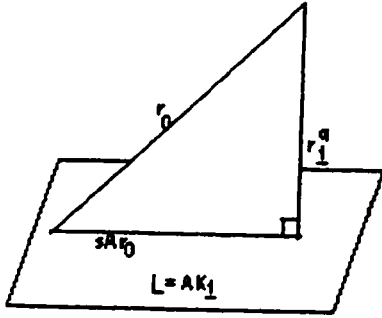


figure 1.5.1.

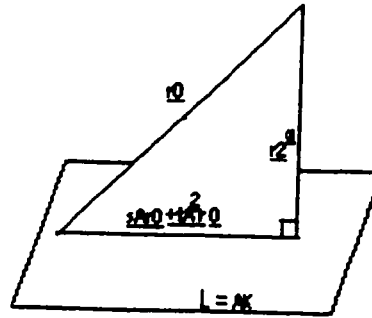


figure 1.5.2.

5.4. Proceed inductively at the  $m^{\text{th}}$  iteration:

Construct  $\mathbf{x}_m^g$  and the corresponding residual,  $\mathbf{r}_m^g$ . These are discovered by defining  $\mathbf{x}_m^g = \mathbf{x}_0 + \mathbf{z}_m^g$ , where  $\mathbf{z}_m^g \in \mathbf{K}_m = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-1}\mathbf{r}_0\}$ , such that  $\mathbf{r}_m^g = \mathbf{b} - \mathbf{A}\mathbf{x}_m^g = \mathbf{r}_0 - \mathbf{A}\mathbf{z}_m^g$  be orthogonal to  $\mathbf{A}\mathbf{K}_m$ . (see figure 1.5.3.)

So,  $\mathbf{r}_m^g \in \mathbf{K}_{m+1} = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^m\mathbf{r}_0\}$ .

The computational method is to select some  $m$  and apply Gram-Schmidt Process to  $\mathbf{K}_m$  to generate orthonormal vectors  $\mathbf{v}_i$ , giving direction of  $\mathbf{r}_{i-1}^g$  (note that this is not the direction of  $\mathbf{r}_{i-1}^g$ ). The procedure is as follows:

Let  $\mathbf{x}_0$  be the initial guess of  $\mathbf{x}$ , then  $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$  will be the initial residual.

Let  $\beta = \|\mathbf{r}_0\|$ , and  $\mathbf{v}_1 = \mathbf{r}_0 / \beta$ . Proceed inductively.

Suppose we have  $\mathbf{K}_{m-1} = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-2}\mathbf{r}_0\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_{m-1}\}$ ,

where  $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_{m-1}\}$  is an orthonormal basis for  $\mathbf{K}_{m-1}$ .

Let  $h_{ij} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$ .

$$\text{Define } \mathbf{u}_m = \mathbf{A}\mathbf{v}_{m-1} - \sum_{i=1}^{m-1} h_{i,m-1} \mathbf{v}_i. \quad (1)$$

$$\text{Then, as before, } \mathbf{v}_{m+1} = \mathbf{u}_{m+1} / \|\mathbf{u}_{m+1}\| = \mathbf{u}_{m+1} / h_{m+1,m}. \quad (\text{figure 1.5.4.}) \quad (2)$$

$$\text{So, } \mathbf{u}_{m+1} = h_{m+1,m} \mathbf{v}_{m+1}.$$

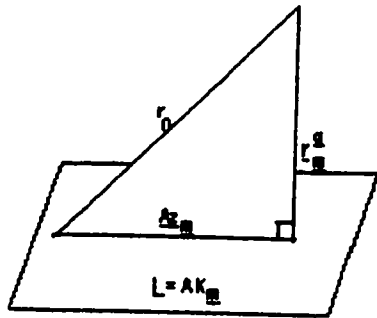


figure 1.5.3 the GMRES method

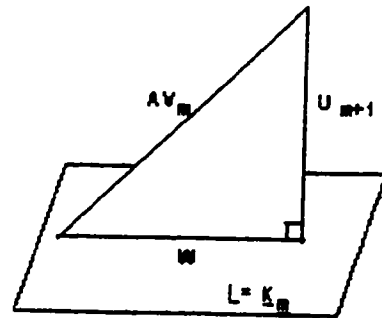


figure 1.5.4 where  $w = \sum_{i=1}^m h_{i,m} \mathbf{v}_i$ .

Define  $\mathbf{H}_m$  to be the  $m \times m$  (Hessenberg) matrix whose entries are the coefficients  $h_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$ ,  $1 \leq i \leq \min\{j+1, m\}$ ,  $1 \leq j \leq m$ , and zero elsewhere.

Let  $\mathbf{V}_m = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m]$ , where  $\mathbf{v}_i$  are column vectors representing  $\mathbf{V}_i$ .

Then it is readily seen that  $\mathbf{H}_m = \mathbf{V}_m^T \mathbf{A} \mathbf{V}_m$ . (3)

The method for finding  $\mathbf{x}_m^8$  is different from Arnoldi's method and is based on the following:

We show that  $\mathbf{A} \mathbf{V}_m = \mathbf{V}_m \mathbf{H}_m + \mathbf{u}_{m+1} \mathbf{e}_m^\top$ .

*Proof:* Since  $\mathbf{u}_{m+1} = \mathbf{A} \mathbf{v}_m - \sum_{i=1}^m h_{i,m} \mathbf{v}_i$ , and since  $\mathbf{H}_m$  is a Hessenberg matrix, and using

equation (1),

$$\mathbf{A} \mathbf{V}_m - \mathbf{V}_m \mathbf{H}_m$$

$$= \mathbf{A} \mathbf{V}_m - \left[ \sum_{i=1}^2 h_{i,1} \mathbf{v}_i, \sum_{i=1}^3 h_{i,2} \mathbf{v}_i, \sum_{i=1}^4 h_{i,3} \mathbf{v}_i, \dots, \sum_{i=1}^m h_{i,m-1} \mathbf{v}_i, \sum_{i=1}^m h_{i,m} \mathbf{v}_i \right]$$

$$= \mathbf{A} \mathbf{V}_m -$$

$$\left[ \left( \sum_{i=1}^1 h_{i,1} \mathbf{v}_i \right) + h_{2,1} \mathbf{v}_2, \left( \sum_{i=1}^2 h_{i,2} \mathbf{v}_i \right) + h_{3,2} \mathbf{v}_3, \dots, \left( \sum_{i=1}^{m-1} h_{i,m-1} \mathbf{v}_i \right) + h_{m,m-1} \mathbf{v}_m, \left( \sum_{i=1}^m h_{i,m} \mathbf{v}_i \right) \right]$$

$$= \mathbf{A} \mathbf{V}_m - \left[ \left( \sum_{i=1}^1 h_{i,1} \mathbf{v}_i \right) + \mathbf{u}_2, \left( \sum_{i=1}^2 h_{i,2} \mathbf{v}_i \right) + \mathbf{u}_3, \dots, \left( \sum_{i=1}^{m-1} h_{i,m-1} \mathbf{v}_i \right) + \mathbf{u}_m, \left( \sum_{i=1}^m h_{i,m} \mathbf{v}_i \right) \right]$$

$$= \mathbf{A} \mathbf{V}_m - \left[ \sum_{i=1}^1 h_{i,1} \mathbf{v}_i, \sum_{i=1}^2 h_{i,2} \mathbf{v}_i, \dots, \sum_{i=1}^{m-1} h_{i,m-1} \mathbf{v}_i, \sum_{i=1}^m h_{i,m} \mathbf{v}_i \right] -$$

$$[\mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_{m-1}, \mathbf{u}_m, \mathbf{0}]$$

$$= \left[ \mathbf{A} \mathbf{v}_1 - \sum_{i=1}^1 h_{i,1} \mathbf{v}_i, \mathbf{A} \mathbf{v}_2 - \sum_{i=1}^2 h_{i,2} \mathbf{v}_i, \dots, \mathbf{A} \mathbf{v}_{m-1} - \sum_{i=1}^{m-1} h_{i,m-1} \mathbf{v}_i, \mathbf{A} \mathbf{v}_m - \sum_{i=1}^m h_{i,m} \mathbf{v}_i \right] -$$

$$[\mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_{m-1}, \mathbf{u}_m, \mathbf{0}]$$

$$= [\mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_{m-1}, \mathbf{u}_m, \mathbf{u}_{m+1}] - [\mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_{m-1}, \mathbf{u}_m, \mathbf{0}]$$

$$= [\mathbf{0}, \mathbf{0}, \dots, \mathbf{0}, \mathbf{0}, \mathbf{u}_{m+1}] = [\mathbf{u}_{m+1}] \times [0, 0, \dots, 0, 1]_{1 \times m}$$

Let  $\mathbf{e}_m^\top = [0, 0, \dots, 0, 1]_{1 \times m}$ , then  $\mathbf{A} \mathbf{V}_m - \mathbf{V}_m \mathbf{H}_m = \mathbf{u}_{m+1} \mathbf{e}_m^\top$ .

Hence,  $\mathbf{A} \mathbf{V}_m = \mathbf{V}_m \mathbf{H}_m + \mathbf{u}_{m+1} \mathbf{e}_m^\top$ . (5)



We show  $\mathbf{A}\mathbf{V}_m = \mathbf{V}_{m+1} \mathbf{H}_m^g$ , where  $\mathbf{H}_m^g$  to be the  $(m+1) \times m$  (Hessenberg) matrix whose nonzero entries are the coefficients  $h_{i,j}$ ,  $1 \leq i \leq j+1$ ,  $1 \leq j \leq m$ , the same as  $\mathbf{H}_m$ , with a bottom row of 0's appended except for  $h_{m+1,m} = (\mathbf{A}\mathbf{v}_m, \mathbf{v}_{m+1})$  in the last column. This Hessenberg matrix,  $\mathbf{H}_m^g$ , is shown in appendix A.

Then, by (5),  $\mathbf{A}\mathbf{V}_m = \mathbf{V}_m \mathbf{H}_m + \mathbf{u}_{m+1} \mathbf{e}_m^T$

$$\begin{aligned} &= \mathbf{V}_m \mathbf{H}_m + h_{m+1,m} \mathbf{v}_{m+1} \mathbf{e}_m^T \\ &= \mathbf{V}_{m+1} \mathbf{H}_m^g \end{aligned} \quad (6)$$

Suppose  $\mathbf{x}_m^g = \mathbf{x}_0 + \mathbf{z}_m^g$  where  $\mathbf{z}_m^g \in \mathbf{K}_m$ , so it can be represented as

$$\mathbf{x}_m^g = \mathbf{x}_0 + \mathbf{V}_m \mathbf{y}_m^g, \quad (7)$$

where  $\mathbf{y}_m^g$  is a  $m \times 1$  vector. We use QR factorization to find a suitable  $\mathbf{z}_m^g$  and  $\mathbf{y}_m^g$  such

that 
$$\|\mathbf{r}_m^g\| = \min\{\|\mathbf{r}_0 - \mathbf{A} \mathbf{z}_m\|\}, \text{ for } \mathbf{z}_m \in \mathbf{K}_m \quad (8)$$

We show that, in fact, 
$$\|\mathbf{r}_m^g\| = \min\{\|\beta \mathbf{e}_{m+1} - \mathbf{H}_m \mathbf{y}_m\|\}, \text{ for } \mathbf{y}_m \in \mathfrak{R}^m. \quad (9)$$

*Proof:* Since  $\mathbf{r}_m^g = (\mathbf{r}_0 - \mathbf{A} \mathbf{z}_m) \perp \mathbf{A} \mathbf{K}_m$ ,  $\mathbf{r}_m^g$  is the vector that minimizes  $\{\|\mathbf{r}_0 - \mathbf{A} \mathbf{z}_m\|\}$ .

So, 
$$\|\mathbf{r}_m^g\| = \min\{\|\mathbf{r}_0 - \mathbf{A} \mathbf{z}_m\|\}, \text{ for } \mathbf{z}_m \in \mathbf{K}_m.$$

Since 
$$\|\mathbf{r}_m^g\| = \|\mathbf{b} - \mathbf{A} \mathbf{x}_m^g\| = \|\mathbf{b} - \mathbf{A} \mathbf{x}_0 - \mathbf{A} \mathbf{z}_m^g\| = \|\mathbf{r}_0 - \mathbf{A} \mathbf{z}_m^g\| = \|\beta \mathbf{v}_1 - \mathbf{A} \mathbf{V}_m \mathbf{y}_m^g\|$$

$$= \|\beta \mathbf{v}_1 - \mathbf{V}_{m+1} \mathbf{H}_m^g \mathbf{y}_m^g\| = \|\beta \mathbf{V}_{m+1} \mathbf{e}_{m+1} - \mathbf{V}_{m+1} \mathbf{H}_m^g \mathbf{y}_m^g\|$$

$$= \|\mathbf{V}_{m+1} (\beta \mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m^g)\| = \|\beta \mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m^g\|,$$

$$\|\mathbf{r}_m^g\| = \min\{\|\beta \mathbf{e}_{m+1} - \mathbf{H}_m \mathbf{y}_m\|\}, \text{ for } \mathbf{y}_m \in \mathfrak{R}^m.$$

A classic way of solving such minimization problems is to factor  $\mathbf{H}_m^g$  into  $\mathbf{Q}_m \mathbf{R}_m$  using plane rotations (where  $\mathbf{Q}$  is orthogonal and  $\mathbf{R}$  is an upper triangular matrix) and find the least-squares solution. We denote by  $\mathbf{y}_m^g$  the solution of the minimization

problem  $\min\{\|\beta\mathbf{e}_{m+1} - \mathbf{H}_m\mathbf{y}_m\|\}$ , for  $\mathbf{y}_m \in \mathbf{R}^m$ . Then the optimal  $\mathbf{x}$  is given by

$$\mathbf{x}_m^g = \mathbf{x}_0 + \mathbf{V}_m \mathbf{y}_m^g \quad [2].$$

The proofs of the following theorems will not be presented since they can be found in most numerical linear algebra textbooks.

**Theorem 1.** The **QR** Factorization: [5]

If  $\mathbf{H}_m^g$  is an  $m \times n$  matrix with linearly independent columns, then  $\mathbf{H}_m^g$  may be factored as  $\mathbf{H}_m^g = \mathbf{QR}$ , where  $\mathbf{Q}$  is an  $m \times n$  matrix whose columns form an orthonormal basis for  $\text{Col } \mathbf{H}_m^g$  and  $\mathbf{R}$  is an  $n \times n$  upper triangular invertible matrix with positive entries on its diagonal.

Least-Squares Solutions:

DEFINITION: If  $\mathbf{H}_m^g$  is  $m \times n$  and  $\mathbf{b}$  is in  $\mathcal{R}^m$ , a least-squares solution of  $\mathbf{H}_m^g \mathbf{x} = \mathbf{b}$  is an

$$\mathbf{x}' \text{ in } \mathcal{R}^n \text{ such that } \|\mathbf{b} - \mathbf{H}_m^g \mathbf{x}'\| \leq \|\mathbf{b} - \mathbf{H}_m^g \mathbf{x}\| \text{ for all } \mathbf{x} \text{ in } \mathcal{R}^n .$$

**Theorem 2.** Given an  $m \times n$  matrix  $\mathbf{H}_m^g$  with linearly independent columns, let  $\mathbf{H}_m^g =$

$\mathbf{QR}$  be a **QR** factorization of  $\mathbf{H}_m^g$  as in the theorem 1. Then for each  $\mathbf{b}$  in

$\mathcal{R}^m$ , the equation  $\mathbf{H}_m^g \mathbf{x} = \mathbf{b}$  has a unique least-squares solution, given by

$$\mathbf{x}' = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}.$$

This method is quite simple to implement because of the special structure of  $\mathbf{H}_m^g$ .

We denote by  $\mathbf{y}_m^g$  the solution of the minimization problem  $\min\{\|\beta\mathbf{e}_{m+1} - \mathbf{H}_m\mathbf{y}_m\|\}$ ,

for  $\mathbf{y}_m \in \mathbf{R}^m$ . Then the optimal  $\mathbf{x}$  is given by  $\mathbf{x}_m^g = \mathbf{x}_0 + \mathbf{V}_m \mathbf{y}_m^g$  [2].

$$\text{Since, by equation (8), } \|\mathbf{r}_m^g\| = \min\{\|\mathbf{r}_0 - \mathbf{A} \mathbf{z}_m\| \text{ for } \mathbf{z}_m \in \mathbf{K}_m\} \dots \dots \dots (8)$$

and since  $\mathbf{K}_1 \subseteq \mathbf{K}_2 \subseteq \mathbf{K}_3 \subseteq \dots \subseteq \mathbf{K}_m \subseteq \dots \subseteq \mathbf{K}_{n-1} \subseteq \mathbf{K}_n$ ;

$$\|r_1^g\| \geq \|r_2^g\| \geq \|r_3^g\| \geq \dots \geq \|r_m^g\| \geq \dots \geq \|r_{n-1}^g\| \geq \|r_n^g\| \dots \dots \dots (10)$$

However, it is desirable to be able to update the factorization of  $H_m^g$  progressively as each column appears, i.e. at every step of the Arnoldi process. This is important because it enables us to obtain the residual norm of  $\|r_m\|$  using approximate solution,  $x_m$ , without computing  $x_{m-1}$  thus allowing us to decide when to stop the process without wasting needless operations[1].

As the number of iterations in GMRES increase, the amount of computations per step also increases. In addition, it sometimes happens that the approximate solution,  $x_m$ , still has a big residual,  $r_m$ . For these reasons, we normally try to restart GMRES after some number of iterations. We will call the number  $m$  here. (Note: We call the number  $m$  “iteration number”; we stop or restart the process at the  $m^{\text{th}}$  iteration.) To restart GMRES, we use the previous approximation  $x_m^{\text{old}}$  as our new initial guess  $x_m^{\text{new}}$  and obtain a new  $V_m$  (or new  $H_m$ ). Then we can apply GMRES again and again until we find an approximate solution which has as small a residual as we expect without increasing the “iteration number”  $m$ . We will discuss more about “restart GMRES” and illustrate this effect in our example in section 1.9.

To summarize, the GMRES method is as follows: [2]

First step, use the Gram-Schmidt process:

For a initial guess  $x_0$ , form  $r_0 = b - Ax_0$ .

Compute  $\beta = \|r_0\|$ , and let  $v_1 = r_0 / \beta$ .

For  $j = 1, 2, \dots, m$  do:

(a) Form  $\mathbf{A}\mathbf{v}_j$  and orthogonalize it against the previous  $\mathbf{v}_1, \dots, \mathbf{v}_j$  via

$$\mathbf{h}_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i), \quad i = 1, 2, \dots, j,$$

$$\mathbf{u}_{j+1} = \mathbf{A}\mathbf{v}_j - \sum_{i=1..j} \mathbf{h}_{i,j} \mathbf{v}_i.$$

$$\mathbf{h}_{j+1,j} = \|\mathbf{u}_{j+1}\|, \text{ and}$$

$$\mathbf{v}_{j+1} = \mathbf{u}_{j+1} / \mathbf{h}_{j+1,j}.$$

(b) Compute the residual norm  $\rho_j = \|\mathbf{b} - \mathbf{A}\mathbf{x}_j\|$  of the solution  $\mathbf{x}_j$  that would be obtained if we stopped at this step.

(c) If  $\rho_j \leq \epsilon$  set  $m = j$  and go to the second step.

Second step, form the approximate solution:

Define  $\mathbf{H}_m^g$  to be the  $(m+1) \times m$  (Hessenberg) matrix whose nonzero entries are the coefficients  $\mathbf{h}_{i,j}$ ,  $1 \leq i \leq j+1$ ,  $1 \leq j \leq m$ , the same as  $\mathbf{H}_m$ , with a bottom row of 0's appended except for  $\mathbf{h}_{m+1,m} = (\mathbf{A}\mathbf{v}_m, \mathbf{v}_{m+1})$  in the last column.

And define  $\mathbf{V}_m = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m]$ .

(a) Find the vector  $\mathbf{y}_m^g$  which minimizes  $\|\beta \mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m\|$  over all vectors  $\mathbf{y}$  in  $\mathbf{R}^m$ .

(b) Compute  $\mathbf{x}_m^g = \mathbf{x}_0 + \mathbf{V}_m \mathbf{y}_m^g$ .

Note: We use a GMRES program which was published in [17].

## Section 1.6 Some conclusions of the case when $\mathbf{H}_m$ becomes singular

Since  $\mathbf{x}_m^A = \mathbf{x}_0 + \beta \mathbf{V}_m \mathbf{H}_m^{-1} \mathbf{e}_m$ , we will not be able to get a solution when  $\mathbf{H}_m$  is singular. Hence, we want to state three theoretical results from [2], in the case that  $\mathbf{H}_m$  is singular. (see [2] for details and proofs )

**Theorem 1.** Suppose  $m$  steps of the Arnoldi process have been taken, and assume that  $\mathbf{H}_m$  is singular. Then

$$\begin{aligned} & \min \{ \|\beta \mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m\| \}, \text{ for } \mathbf{y}_m^g \in \mathbf{R}^m \\ & = \min \{ \|\beta \mathbf{e}_{m+1} - \mathbf{H}_{m-1}^g \mathbf{y}_{m-1}\| \}, \text{ for } \mathbf{y}_{m-1}^g \in \mathbf{R}^{m-1}. \end{aligned} \quad (\text{a})$$

If  $\mathbf{y}_j^g$  is the solution of  $\min \{ \|\beta \mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m\| \}$  with  $m$  replaced by  $j$ ,  $j = m$  or  $m-1$ , then  $\mathbf{y}_m^g = ((\mathbf{y}_{m-1}^g)^T, 0)^T$ , and it follows that  $\mathbf{x}_m^g = \mathbf{x}_{m-1}^g$ .

Conversely, suppose  $m$  steps of the Arnoldi process have been taken and that (a) holds, Then  $\mathbf{H}_m$  is singular [2].

This theorem gives a clear indication of how the performance of the Arnoldi and GMRES algorithms are interrelated. If  $\mathbf{H}_m$  is singular, so that  $\mathbf{x}_m^A$  does not exist, then GMRES is unsuccessful in reducing the norm of the residual, and  $\mathbf{x}_m^g = \mathbf{x}_{m-1}^g$ .

**Definition:** Let us define the function  $g(x)$  by  $g(x) = \|(Ax - b)\|^2$ .

We say that the vector  $v$  is an ascent direction for  $g$  at  $x$  if

$$g(x + \lambda v) > g(x) \text{ for all nonzero } \lambda.$$

**Theorem 2.** Suppose that  $m$  steps of the Arnoldi process have been taken. Then  $\mathbf{H}_m$  is singular if and only if  $\mathbf{v}_m$  is an ascent direction for  $g$  at  $\mathbf{x}_{m-1}^g$ .

Theorem 1 and theorem 2 indicate what happens when the Arnoldi and GMRES algorithms are performing poorly. If  $\mathbf{H}_m$  is nearly singular, or equivalently, if  $\|\mathbf{r}_m^g\|$  is only slightly smaller than  $\|\mathbf{r}_{m-1}^g\|$ , then the new Krylov vector  $\mathbf{v}_m$  is close to being an ascent direction for  $g$  at  $\mathbf{x}_{m-1}^g$  [2].

**Theorem 3.** Suppose  $m$  steps of the Arnoldi process have been taken, and assume that

$\mathbf{H}_m$  is singular. Then

$$\begin{aligned} & \min\{ \|\mathbf{V}_m^T (\mathbf{b} - \mathbf{A}\mathbf{x}_m)\| \}, \text{ for } \mathbf{x}_m \in \mathbf{x}_0 + \mathbf{K}_m \\ & = \min\{ \|\beta\mathbf{e}_m - \mathbf{H}_m^A \mathbf{y}_m\| \}, \text{ for } \mathbf{y}_m \in \mathbf{R}^m \end{aligned} \tag{b}$$

$$= \min\{ \|\beta\mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m\| \}, \text{ for } \mathbf{y}_m \in \mathbf{R}^m \tag{c}$$

Furthermore,  $\mathbf{y}_m^{ls} = \mathbf{y}_m^g$ , where  $\mathbf{y}_m^{ls}$  and  $\mathbf{y}_m^g$  are the solutions to the last two problems (b) and (c) (respectively) [2].

Note: Detailed proofs for these three theorems are provided in [2], based on the condition of  $\mathbf{v}_m \neq 0$ .

## Section 1.7 Cases that make $\mathbf{H}_m$ become singular

Let  $\mathbf{H}_m(:, j)$  denote the  $j^{\text{th}}$  column of  $\mathbf{H}_m$  and  $\mathbf{H}_m(i, :)$  denote the  $i^{\text{th}}$  row of  $\mathbf{H}_m$ .

Suppose that  $\mathbf{H}_i$  is nonsingular and  $\mathbf{H}_m$  is singular while  $\mathbf{v}_i \neq \mathbf{0}$  for  $i = 1, 2, 3, \dots, m-1$ .

Here, we want to describe the behavior of GMRES when  $\mathbf{H}_m$  becomes singular, from the view of geometry. There are four possibilities.

- (a) The  $m^{\text{th}}$  column of  $\mathbf{H}_m$  is a zero vector.
- (b) The entries of the  $m^{\text{th}}$  row of  $\mathbf{H}_m$  are all zeros.
- (c) The  $m^{\text{th}}$  column of  $\mathbf{H}_m$  is a linear combination of the first to the  $(m-1)^{\text{th}}$  columns of  $\mathbf{H}_m$ .
- (d) The  $m^{\text{th}}$  row of  $\mathbf{H}_m$  is a linear combination of the first to the  $(m-1)^{\text{th}}$  rows of  $\mathbf{H}_m$ .

We can further discuss (a) and (b):

Since  $h_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$ , we have:

- (a) If the  $m^{\text{th}}$  column of  $\mathbf{H}_m$  is the zero vector then either (i)  $\mathbf{A}\mathbf{v}_m = \mathbf{0}$  or (ii)  $\mathbf{A}\mathbf{v}_m \perp \mathbf{v}_i$  for  $i = 1, 2, \dots, m$  where  $\mathbf{A}\mathbf{v}_m \neq \mathbf{0}$ .

*Proof:* Since  $h_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$ ,  $h_{i,m} = (\mathbf{A}\mathbf{v}_m, \mathbf{v}_i)$  for  $i = 1, 2, 3, \dots, m$ .

For  $i = 1, 2, 3, \dots, m$ , since  $h_{i,m} = 0$ ,  $(\mathbf{A}\mathbf{v}_m, \mathbf{v}_i) = 0$ .

Hence, either  $\mathbf{A}\mathbf{v}_m = \mathbf{0}$  or  $\mathbf{A}\mathbf{v}_m \perp \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_m\}$ .

In other words, either  $\mathbf{A}\mathbf{v}_m = \mathbf{0}$  or

$\mathbf{v}_{m+1}$  can be  $\mathbf{A}\mathbf{v}_m / \|\mathbf{A}\mathbf{v}_m\|$  if  $\mathbf{A}\mathbf{v}_m \neq \mathbf{0}$ .

- (b) If the entries of  $m^{\text{th}}$  row of  $\mathbf{H}_m$  are all zeros then (i)  $\mathbf{A}\mathbf{v}_m = \mathbf{0}$ .

*Proof:* If the entries of  $m^{\text{th}}$  row of  $\mathbf{H}_m$  are all zeros then

$$(\mathbf{A}\mathbf{v}_{m-1}, \mathbf{v}_m) = 0 \text{ and } (\mathbf{A}\mathbf{v}_m, \mathbf{v}_m) = 0.$$

Suppose  $\mathbf{v}_m \neq 0$ .

Since  $\mathbf{v}_m = \varepsilon \mathbf{A}\mathbf{v}_{m-1} + \sum_{i=1}^{m-1} \delta_i \mathbf{v}_i$ , and since  $\mathbf{v}_m \perp \text{span} \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_{m-1}\}$ ,

$$1 = (\mathbf{v}_m, \mathbf{v}_m) = (\mathbf{v}_m, \varepsilon \mathbf{A}\mathbf{v}_{m-1}) + \sum_{i=1}^{m-1} \delta_i (\mathbf{v}_m, \mathbf{v}_i) = 0 \text{ (since } (\mathbf{A}\mathbf{v}_{m-1}, \mathbf{v}_m) = 0).$$

This is impossible to happen.

Hence,  $\mathbf{v}_m = 0$ . So,  $\mathbf{A}\mathbf{v}_m = 0$ .

Hence, we conclude that if (a), (b), (c) or (d) hold, then

(i)  $\mathbf{A}\mathbf{v}_m = 0$  ( $\mathbf{v}_m = 0$ ), or

(ii)  $\mathbf{A}\mathbf{v}_m \perp \mathbf{v}_i$  for  $i = 1, 2, \dots, m$  where  $\mathbf{A}\mathbf{v}_m \neq 0$ , or

(iii)  $\mathbf{H}_m(:, m) = \sum_{j=1}^{m-1} a_j \mathbf{H}_m(:, j)$ , or

(iv)  $\mathbf{H}_m(m, :) = \sum_{j=1}^{m-1} a_j \mathbf{H}_m(j, :)$  hold.

Therefore, if  $\mathbf{H}_m$  is singular, then (i), or (ii), or (iii), or (iv) hold. Next, in **section 1.8**, we prove that if (i), or (ii), or (iii) hold, then  $\mathbf{x}_{m-1}^g = \mathbf{x}_m^g$ .

Note: The proof for: “If (iv) holds, then  $\mathbf{x}_{m-1}^g = \mathbf{x}_m^g$ ” will not be provided in this paper.

The reader can find the proof from [2]: If  $\mathbf{H}_m$  is singular, then  $\mathbf{x}_{m-1}^g = \mathbf{x}_m^g$  (when  $\mathbf{v}_m \neq 0$ ).



## Section 1.8 Proofs that $\mathbf{x}_{m-1}^g = \mathbf{x}_m^g$ in some cases

Before the proofs, we have to know :

(a) When we apply GMRES to find the residual  $\mathbf{r}$ , there are only two cases that will

occur: either  $\|\mathbf{r}_{m-1}^g\| < \|\mathbf{r}_m^g\|$  or  $\|\mathbf{r}_{m-1}^g\| = \|\mathbf{r}_m^g\|$  (see 9.1 in section 1.9 for proof).

$$(b) \|\mathbf{r}_m^g\| = \min\{\|\beta\mathbf{e}_{m+1} - \mathbf{H}_m\mathbf{y}_m\|\}, \quad (9)$$

where  $\mathbf{y}_m \in \mathbf{R}^m$  denotes the solution for  $\min\{\|\beta\mathbf{e}_{m+1} - \mathbf{H}_m\mathbf{y}_m\|\}$ , for  $\mathbf{y}_m \in \mathfrak{R}^m$ .

$$(c) \mathbf{x}_m^g = \mathbf{x}_0 + \mathbf{V}_m\mathbf{y}_m^g, \text{ where } \mathbf{V}_m\mathbf{y}_m^g = \mathbf{z}_m^g \in \mathbf{K}_m \text{ gives the solution for } \min\{\|\mathbf{r}_0 - \mathbf{A}\mathbf{z}_m\|\} \quad (7)$$

(d) If  $\|\mathbf{r}_{m-1}^g\| = \|\mathbf{r}_m^g\|$ , then  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$ .

*Proof:* Let  $\mathbf{K}_s$  be the Krylov subspace and  $\mathbf{r}_s^g = \mathbf{r}_0 - \mathbf{A}\mathbf{z}_s^g$  for  $\mathbf{z}_s^g \in \mathbf{K}_s$  as were

mentioned in the previous sections. Suppose  $\|\mathbf{r}_{m-1}^g\| = \|\mathbf{r}_m^g\|$ .

Since, by equation (8),  $\|\mathbf{r}_s^g\| = \min\{\|\mathbf{r}_0 - \mathbf{A}\mathbf{z}_s\|\}$  for  $\mathbf{z}_s \in \mathbf{K}_s$ ,

$\mathbf{z}_{m-1}^g$  is the least-square solution to  $\|\mathbf{r}_{m-1}^g\| = \min\{\|\mathbf{r}_0 - \mathbf{A}\mathbf{z}_{m-1}\|\}$  in  $\mathbf{K}_{m-1}$  and

$\mathbf{z}_m^g$  is the least-square solution to  $\|\mathbf{r}_m^g\| = \min\{\|\mathbf{r}_0 - \mathbf{A}\mathbf{z}_m\|\}$  in  $\mathbf{K}_m$ .

Since  $\mathbf{K}_{m-1} \subseteq \mathbf{K}_m$ , and since  $\|\mathbf{r}_{m-1}^g\| = \|\mathbf{r}_m^g\|$ , by the uniqueness property

in theorem 2 of section 1.5,  $\mathbf{z}_{m-1}^g = \mathbf{z}_m^g$ . In other words,  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$ .

(e) Also, if  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$ , then we have  $\mathbf{x}_{m-1}^g = \mathbf{x}_m^g$ . This follows since

$$\mathbf{b} - \mathbf{A}\mathbf{x}_{m-1}^g = \mathbf{r}_0 - \mathbf{A}\mathbf{z}_{m-1}^g = \mathbf{r}_{m-1}^g = \mathbf{r}_m^g = \mathbf{r}_0 - \mathbf{A}\mathbf{z}_m^g = \mathbf{b} - \mathbf{A}\mathbf{x}_m^g \text{ and } \mathbf{A} \text{ is non-singular.}$$

So we can conclude that if  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$ , then  $\mathbf{x}_{m-1}^g = \mathbf{x}_m^g$ .

In the next three proofs, we prove:  $\|\mathbf{r}_{m-1}^g\| = \|\mathbf{r}_m^g\|$  or  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$ , in the case of (i), or

(ii), or (iii) hold; and then  $\mathbf{x}_{m-1}^g = \mathbf{x}_m^g$  follows immediately.

**8.1** If  $\mathbf{A}\mathbf{v}_m = \mathbf{0}$ , then  $\mathbf{x}_{m-1}^g = \mathbf{x}_{m-1}^A = \mathbf{x}$  and therefore  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$

Since we assumed that  $\mathbf{A}$  is nonsingular,  $\mathbf{A}\mathbf{v}_m = \mathbf{0}$  happens if and only if  $\mathbf{v}_m = \mathbf{0}$ .

Thus Gram-Schmidt failed to construct an additional orthonormal vector, so  $\mathbf{A}\mathbf{v}_{m-1} \in \mathbf{K}_{m-1}$ , and then  $\mathbf{K}_{m-1} = \mathbf{K}_m = \mathbf{K}_{m+1} = \dots = \mathbf{K}_n$ . This also implies that  $\mathbf{r}_0 \in \mathbf{A}\mathbf{K}_{m-1}$  (lemma 1.) and therefore  $\mathbf{r}_m = \mathbf{0}$  (lemma 2.). This situation has often been referred to as a “happy break down” and we have found the exact solution [2].

Note: Since  $0 \leq \|\mathbf{r}_m^g\| \leq \|\mathbf{r}_{m-1}^g\| = 0$ ,  $\|\mathbf{r}_m^g\| = 0$ . i.e.  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g = \mathbf{0}$ .

**Lemma 1.**  $\mathbf{K}_{m-1} = \mathbf{K}_m$  if and only if  $\mathbf{r}_0 \in \mathbf{A}\mathbf{K}_{m-1}$ .

*Proof:* Suppose  $\mathbf{K}_{m-1} = \mathbf{K}_m$  and  $\mathbf{A}$  is non-singular. This occurs

if and only if there exists some  $a_i$  such that  $\mathbf{A}^{m-1}\mathbf{r}_0 = \sum_{i=0}^{m-2} a_i \mathbf{A}^i \mathbf{r}_0$ . Let  $a_s$  be the

first non-zero  $a_i$  (so  $a_k = 0$  for  $k = 1, 2, 3, \dots, s-1$ ), then

$$\mathbf{A}^{m-1}\mathbf{r}_0 = \sum_{i=0}^{m-2} a_i \mathbf{A}^i \mathbf{r}_0$$

if and only if  $a_s \mathbf{A}^s \mathbf{r}_0 = \mathbf{A}^{m-1} \mathbf{r}_0 - \sum_{i=s+1}^{m-2} a_i \mathbf{A}^i \mathbf{r}_0 = \mathbf{A}^s (\mathbf{A}^{m-s-1} \mathbf{r}_0 - \sum_{i=1}^{m-s-2} a_{i+s} \mathbf{A}^i \mathbf{r}_0)$

if and only if  $a_s \mathbf{r}_0 = \mathbf{A}^{m-s-1} \mathbf{r}_0 - \sum_{i=1}^{m-s-2} a_{i+s} \mathbf{A}^i \mathbf{r}_0$  where  $a_s \neq 0$

if and only if  $a_s \mathbf{r}_0 = \mathbf{A} (\mathbf{A}^{m-s-2} \mathbf{r}_0 - \sum_{i=0}^{m-s-3} a_{i+s+1} \mathbf{A}^i \mathbf{r}_0)$  where  $a_s \neq 0$

if and only if  $\mathbf{r}_0 \in \mathbf{A}\mathbf{K}_{m-s-1} \subset \mathbf{A}\mathbf{K}_{m-1}$ .

**Lemma 2.**  $\mathbf{r}_0 \in \mathbf{AK}_{m-1}$  if and only if  $\mathbf{x}_{m-1}^{\mathcal{B}} = \mathbf{x}_{m-1}^{\mathcal{A}} = \mathbf{x}$ , the exact solution.

*Proof.* (a) Suppose  $\mathbf{r}_0 \in \mathbf{AK}_{m-1}$ , then  $\mathbf{r}_0 = \mathbf{Az} \in \mathbf{AK}_{m-1}$ .

So  $\|\mathbf{r}_0 - \mathbf{Az}_m\| = 0$  for  $\mathbf{z} \in \mathbf{K}_{m-1}$ .

But  $\|\mathbf{r}_{m-1}^{\mathcal{B}}\|$  equals the minimum for such expressions, then  $\mathbf{z}_{m-1}^{\mathcal{B}} = \mathbf{z}$ .

Hence,  $\mathbf{r}_{m-1}^{\mathcal{B}} = \mathbf{r}_0 - \mathbf{Az}_{m-1}^{\mathcal{B}} = \mathbf{r}_0 - \mathbf{Az} = \mathbf{0}$ .

Therefore,  $\mathbf{0} = \mathbf{r}_{m-1}^{\mathcal{B}} = \mathbf{b} - \mathbf{Ax}_0 - \mathbf{Az}_{m-1}^{\mathcal{B}} = \mathbf{b} - \mathbf{Ax}_{m-1}^{\mathcal{B}}$ .

In other words,  $\mathbf{x}_{m-1}^{\mathcal{B}} = \mathbf{x}$ .

Also since  $\mathbf{z}_{m-1}^{\mathcal{A}}$  is the unique vector in  $\mathbf{K}_{m-1}$  such that

$\mathbf{r}_{m-1}^{\mathcal{A}} = \mathbf{r}_0 - \mathbf{Az}_{m-1}^{\mathcal{A}} \perp \mathbf{K}_{m-1}$  and since  $\mathbf{r}_0 - \mathbf{Az} = \mathbf{0} \perp \mathbf{K}_{m-1}$ ,  $\mathbf{z}_{m-1}^{\mathcal{A}} = \mathbf{z}$ .

Hence,  $\mathbf{r}_{m-1}^{\mathcal{A}} = \mathbf{r}_0 - \mathbf{Az}_{m-1}^{\mathcal{A}} = \mathbf{r}_0 - \mathbf{Az} = \mathbf{0}$ .

Therefore,  $\mathbf{0} = \mathbf{r}_{m-1}^{\mathcal{A}} = \mathbf{b} - \mathbf{Ax}_0 - \mathbf{Az}_{m-1}^{\mathcal{A}} = \mathbf{b} - \mathbf{Ax}_{m-1}^{\mathcal{A}}$ .

In other words,  $\mathbf{x}_{m-1}^{\mathcal{A}} = \mathbf{x}$ . So,  $\mathbf{x}_{m-1}^{\mathcal{B}} = \mathbf{x}_{m-1}^{\mathcal{A}} = \mathbf{x}$ .

(b) Suppose that  $\mathbf{x}_{m-1}^{\mathcal{B}} = \mathbf{x}$ , then  $\mathbf{x}_{m-1}^{\mathcal{B}} = \mathbf{x}_0 + \mathbf{z}_{m-1}^{\mathcal{B}}$ .

So,  $\mathbf{0} = \mathbf{b} - \mathbf{Ax}_{m-1}^{\mathcal{B}} = \mathbf{b} - \mathbf{A}(\mathbf{x}_0 + \mathbf{z}_{m-1}^{\mathcal{B}}) = \mathbf{b} - \mathbf{Ax}_0 - \mathbf{Az}_{m-1}^{\mathcal{B}} = \mathbf{r}_0 - \mathbf{Az}_{m-1}^{\mathcal{B}}$ .

Hence,  $\mathbf{r}_0 = \mathbf{Az}_{m-1}^{\mathcal{B}} \in \mathbf{AK}_{m-1}$ .

Therefore, we have the following corollaries:

**Corollary 1.**  $\mathbf{A}^{m-1} \mathbf{r}_0 \in \mathbf{K}_{m-1}$  if and only if  $\mathbf{x}_{m-1}^{\mathcal{B}} = \mathbf{x}_{m-1}^{\mathcal{A}} = \mathbf{x}$ , the exact solution.

*Proof:*  $\mathbf{A}^{m-1} \mathbf{r}_0 \in \mathbf{K}_{m-1} = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-2}\mathbf{r}_0\}$

if and only if  $\mathbf{K}_{m-1} = \text{span}\{\mathbf{r}_0, \mathbf{Ar}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{m-2}\mathbf{r}_0, \mathbf{A}^{m-1}\mathbf{r}_0\} = \mathbf{K}_m$

if and only if  $\mathbf{r}_0 \in \mathbf{AK}_{m-1}$  (lemma 1.)

if and only if  $\mathbf{x}_{m-1}^{\mathcal{B}} = \mathbf{x}_{m-1}^{\mathcal{A}} = \mathbf{x}$  (lemma 2.).

Note that if  $r_0$  is an eigen-vector of  $A^{m-1}$ , then  $A^{m-1} r_0 \in K_{m-1}$ . So,  $x_{m-1}^g = x_{m-1}^A = x$ .

Corollary 2 follows from corollary 1:

**Corollary 2.** Suppose  $A^{k-1} r_0 \notin K_{k-1}$  for  $k < m$ , then we can never have  $x_s^g = x$  or  $x_s^A = x$  for  $s < m-1$ .

Hence, this “m-1” will be the minimum iteration number we can choose in order to get an exact solution without restarting the GMRES method.

**Corollary 3.** Suppose that  $r_0$  is an eigen- vector of  $A^{m-1}$ , then we have  $x_i^g \neq x$  for  $i = 1, 2, \dots, m-2$ , by corollary 2, and  $x_{m-1}^g = x$  by corollary 1.

*Proof:* Since  $r_0$  is an eigen- vector of  $A^{m-1}$ , there exist  $\lambda$  such that  $A^{m-1} r_0 = \lambda r_0 \in K_{m-1}$ .

Hence, we have  $x_i^g \neq x$  for  $i = 1, 2, \dots, m-2$ , by corollary 2, and  $x_{m-1}^g = x$  by corollary 1.

Some examples are given as follow:

**Example 8.1.1:**  $r_0$  is an eigen-vector of  $A^2$  (then we have  $x_2^g = x$ , by corollary 1 and  $x_1^g \neq x$  by corollary 2.)

$$A = \begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & -1 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & -1 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & -1 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & -1 & 0 \\ 7 & 0 & 0 & 0 & 0 & 0 & -1 \end{matrix}$$

$b =$	$r_0 =$	$Ar_0 =$	$A^2 r_0 = r_0 =$	$x_0 =$
1	1	1	1	0
1	0	2	0	-1.0000
1	0	3	0	-1.0000
1	0	4	0	-1.0000
1	0	5	0	-1.0000
1	0	6	0	-1.0000
1	0	7	0	-1.0000

(a) Restart after  $m = 2$  iterations, and do not do any repetition.  
 (Here we show  $\mathbf{x}_2^g = \mathbf{x}$  without restarting GMRES)

$H_2 =$  (note that  $H_3$  is singular)  
 $\begin{matrix} 11.8322 & -0.9964 \\ 0 & -0.0845 \end{matrix}$

the residuals $e =$	ans $\mathbf{x}_2 = \mathbf{x} =$
1.0e-015 *	1
0.2220	2
0.4441	3
-0.2220	4
0.8882	5
0.2220	6
-0.4441	7
0.6661	

(b) Restart after  $m = 1$  iterations, and do not do any repetition.  
 ( Here we show that  $\mathbf{x}_1^g \neq \mathbf{x}$  )

$H_1 = 11.8322$

the residuals $e =$	ans $\mathbf{x}_1 =$
-0.9929	0.0071
0.0143	-1.0000
0.0214	-1.0000
0.0286	-1.0000
0.0357	-1.0000
0.0429	-1.0000
0.500	-1.0000
0.501	

**Example 8.1.2 :**  $\mathbf{r}_0$  is an eigen- vector of  $A^3$ . (Then we will have  $\mathbf{x}_1^g \neq \mathbf{x}$ , and  $\mathbf{x}_2^g \neq \mathbf{x}$ , by corollary 2 and  $\mathbf{x}_3^g = \mathbf{x}$  by corollary 1.)

$A =$

1	-1	0	0	0	0	0
1	0	0	0	0	0	0
0	1	-1	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	1	0	0
0	0	0	0	0	1	0
0	0	0	0	0	0	1

$b =$	$x_0 =$	$r_0 =$	$r_0 A =$	$r_0 A^2 =$	$r_0 A^3 = -r_0$
1	1	1	1	0	-1
1	1	0	1	1	0
1	0	0	0	1	0
1	1	0	0	0	0
1	1	0	0	0	0
1	1	0	0	0	0
1	1	0	0	0	0

(a) Restart after  $m = 1$  iterations, and do not do any repetition. (Here we show  $x_1^B \neq x$ .)

the residuals $e =$	ans $x_1 =$	$H_1 = 1.4142$
-0.5000	1.5000	
0.5000	1.0000	
0	0	
0	1.0000	
0	1.0000	
0	1.0000	
0	1.0000	

(b) Restart after  $m = 2$  iterations, and do not do any repetition. (Here we show  $x_2^B \neq x$ .)

the residuals $e =$	ans $x_2 =$	$H_2 = 1.4142$	-0.7071
-0.3333	1.3333	0	1.2247
0.3333	0.6667		
-0.3333	0		
0	1.0000		
0	1.0000		
0	1.0000		
0	1.0000		

(c) Restart after  $m = 3$  iterations, and do not do any repetition. (Here we show  $x_3^B = x$ .)

the residuals $e =$	ans $x_3 = x$	$H_3 =$	1.4142	-0.7071	0
0	1	0	0	1.2247	-0.8165
0	0	0	0	0	-0.5774
0	-1	0	0	0	0
0	1				
0	1				
0	1				
0	1				

**8.2.  $\mathbf{Av}_m \perp \mathbf{v}_i$  for  $i = 1, 2, 3, \dots, m$ , where  $\mathbf{Av}_m \neq 0$ . Then  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$ .**

Suppose  $\mathbf{Av}_m \perp \mathbf{v}_i$  for  $i = 1, 2, 3, \dots, m$ , where  $\mathbf{Av}_m \neq 0$ , then  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$ .

*proof:* Suppose that  $\mathbf{Av}_m \perp \mathbf{v}_i$  for  $i = 1, 2, 3, \dots, m$  and  $\mathbf{Av}_m \neq 0$ .

Since  $\mathbf{r}_m^g = \mathbf{r}_0 - \mathbf{Az}_m$ , for  $\mathbf{z}_m \in \mathbf{K}_m$ , there exists  $a_i$  such that  $\mathbf{z}_m = \sum_{i=1}^m a_i \mathbf{v}_i$ .

Since  $\mathbf{H}_{m-1}$  is non-singular and  $\mathbf{Av}_m \neq 0$ ,  $\mathbf{v}_i \neq 0$  for  $i = 1, 2, 3, \dots, m$ .

Suppose that  $a_m \neq 0$  such that  $\mathbf{r}_m^g = \mathbf{r}_0 - \mathbf{A} \left( \sum_{i=1}^m a_i \mathbf{v}_i \right)$ ,

then  $\mathbf{r}_m^g = \mathbf{r}_0 - \left( \sum_{i=1}^m a_i \mathbf{Av}_i \right) = \mathbf{r}_0 - \left( \sum_{i=1}^{m-1} a_i \mathbf{Av}_i \right) - a_m \mathbf{Av}_m$ .

Let  $\mathbf{T}_{m-1} = \mathbf{r}_0 - \left( \sum_{i=1}^{m-1} a_i \mathbf{Av}_i \right)$ , then  $\|\mathbf{T}_{m-1}\| \geq \|\mathbf{r}_{m-1}^g\|$  and  $\mathbf{T}_{m-1} \in \mathbf{K}_m$ .

Since  $\mathbf{Av}_m \perp \mathbf{v}_i$  for  $i = 1, 2, 3, \dots, m$ ,  $\mathbf{Av}_m \perp \mathbf{K}_m$ .

Hence,  $\mathbf{Av}_m \perp \mathbf{T}_{m-1}$ .

Since  $\mathbf{r}_m^g = \mathbf{T}_{m-1} - a_m \mathbf{Av}_m$ , by the Pythagorean theorem (figure 1.8.1),

$\|\mathbf{r}_m^g\| > \|\mathbf{T}_{m-1}\| \geq \|\mathbf{r}_{m-1}^g\|$ . (if  $a_m \mathbf{Av}_m \neq 0$ )

This result contradicts the GMRES's result:  $\|\mathbf{r}_{m-1}^g\| \geq \|\mathbf{r}_m^g\|$ .

Therefore,  $a_m = 0$ .

In other words,  $\mathbf{r}_m^g = \mathbf{r}_0 - \left( \sum_{i=1}^{m-1} h_i \mathbf{Av}_i \right) \in \mathbf{K}_m$ . So,  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g \in \mathbf{K}_m$ .

Furthermore, if  $\mathbf{Av}_s \perp \mathbf{v}_i$  for  $i = 1, 2, 3, \dots, s$  where  $\mathbf{Av}_s \neq 0$  for  $s = m, m+1, \dots, t$ , then we

are not going to have a better solution than  $r_{m-1}^g$  for  $m \leq s \leq t$ .

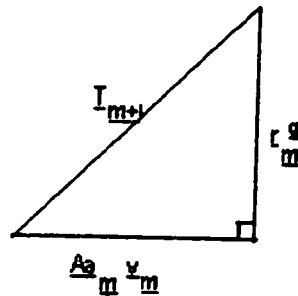


figure 1.8.1

From the following example we show how this effects our discussion:

**Example:** Let  $A$  be the permutation matrix sending  $e_1 \rightarrow e_2 \rightarrow \dots \rightarrow e_n \rightarrow e_1$ ,

where  $e_i$  is the  $i^{\text{th}}$  standard basis vector in  $R_n$ .

Let  $Ax = b$  where  $b = e_1$  and  $x_0 = 0$ . Then  $r_0 = e_1$ .

Therefore,  $v_i = e_i$  and  $Av_i = e_{i+1}$  for  $i = 1, 2, \dots, n$  where  $e_{n+1} = e_1$ .

In other words,

(a) if  $s < n$ , then  $Av_s \perp v_i$  for  $i = 1, 2, 3, \dots, s$ , and  $r_0 = r_s^g = e_1$ .

(Hence,  $H_s$  is singular for  $s < n$ .)

(b) if  $s = n$ , then  $Av_n \perp v_i$  for  $i = 2, 3, \dots, n$ , and  $(Av_n, v_1) = 1$ .

(Hence,  $r_n^g = 0$  and we found the exact solution.)

In data 1 and data 2, we provide some data for the case that  $n = 7$  and vector  $b$  as in the above example.



Data 1:

$$\begin{array}{r}
 A = \\
 \begin{array}{cccccccc}
 0 & 0 & 0 & 0 & 0 & 0 & 1 & \\
 1 & 0 & 0 & 0 & 0 & 0 & 0 & \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 & \\
 0 & 0 & 1 & 0 & 0 & 0 & 0 & \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 & \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 & \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
 \end{array}
 \end{array}
 \qquad
 \begin{array}{r}
 b = \\
 \begin{array}{c}
 1 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{array}
 \end{array}$$

(a) In this example  $H_1, H_2, H_3, \dots,$  and  $H_6$  are all singular because  $Av_m \perp v_i$  for  $i = 1, 2, \dots, m$  by choosing  $x_0 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$ . (So,  $x_0 = x_1^g = x_2^g = x_3^g = x_4^g = x_5^g = x_6^g$ .)

Restart after  $m = 3$  iterations, and do not do any repetition.

$$\begin{array}{r}
 \text{the residuals } e = \\
 \begin{array}{c}
 -1 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{array}
 \end{array}
 \qquad
 \begin{array}{r}
 \text{ans } x_3^g = \\
 \begin{array}{c}
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{array}
 \end{array}$$

$$\begin{array}{r}
 H_3 = \\
 \begin{array}{ccc}
 0 & 0 & 0 \\
 1 & 0 & 0 \\
 0 & 1 & 0
 \end{array}
 \end{array}
 \qquad
 \begin{array}{r}
 H_4 = \\
 \begin{array}{cccc}
 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 \\
 0 & 0 & 1 & 0
 \end{array}
 \end{array}$$

$$\begin{array}{r}
 H_5 = \\
 \begin{array}{ccccc}
 0 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 1 & 0
 \end{array}
 \end{array}
 \qquad
 \begin{array}{r}
 H_6 = \\
 \begin{array}{cccccc}
 0 & 0 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0
 \end{array}
 \end{array}$$

Restart after  $m = 6$  iterations, and do not do any repetition.

$$\begin{array}{r}
 \text{the residuals } e = \\
 \begin{array}{c}
 -1 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{array}
 \end{array}
 \qquad
 \begin{array}{r}
 \text{ans } x_6^g = \\
 \begin{array}{c}
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0
 \end{array}
 \end{array}$$

Since  $\mathbf{A}\mathbf{v}_m \perp \mathbf{v}_i$  for  $i = 1, 2, \dots, m$ , we have  $\mathbf{H}_m$  be singular for  $i = 1, 2, \dots, 6$ .

So, initial guess  $\mathbf{x}_0 = \mathbf{x}_i^g =$  first restart  $\mathbf{x}_0 = \mathbf{x}_i^g =$  second restart  $\mathbf{x}_0 = \mathbf{x}_i^g = \dots$

for  $i = 1, 2, \dots, 6$ . Hence, no matter which  $m$  we choose (unless  $\mathbf{H}_m$  is not singular), the restart method will still give the initial guess as the result. Since  $\mathbf{A}$  is a  $7 \times 7$  matrix,

$\mathbf{K}_7 = \mathbf{K}_8$ . Therefore, by 5.1,  $\mathbf{x}_7 = \mathbf{x}$ . Hence, we have the next example:

(b)  $\mathbf{H}_7$  is non-singular and we can have the exact solution:

Restart after  $m=7$  iterations, and do not do any repetition.

$\mathbf{H}_7 =$	the residuals $\mathbf{e} =$	ans $\mathbf{x}_7 =$
0 0 0 0 0 0 1	0	0
1 0 0 0 0 0 0	0	0
0 1 0 0 0 0 0	0	0
0 0 1 0 0 0 0	0	0
0 0 0 1 0 0 0	0	0
0 0 0 0 1 0 0	0	0
0 0 0 0 0 1 0	0	1

Since we do not expect to have a large size of  $\mathbf{H}_m$ , we want to try other initial guesses,  $\mathbf{x}_0$  (see data 2); because with same initial guess as above example,  $\mathbf{H}_m$  is singular and the result is always the same as our initial guess, unless  $m = n$  (no progress unless  $\mathbf{H}_m$  is not singular). Hence, we want to choose another initial guess in order to prevent this situation from happening again.

Data 2: We want to try two new initial guesses:  $\mathbf{x}_0 = [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7]^T$  and  $\mathbf{x}_0 = \text{random vector} \in \mathcal{R}^7$ .

(a)  $\mathbf{x}_0 = [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7]^T$ :

Restart after  $m=3$  iterations, and do not do any repetition.

the residuals $\mathbf{e} =$	ans $\mathbf{x}_3 =$
0.7691	-4.2211
-4.2211	1.1475
1.1475	0.2527
0.2527	1.4732
1.4732	1.5718
1.5718	1.6705
1.6705	0.7691

Restart after  $m=3$  iterations, repeat this 19 times; (Since  $\mathbf{x}_0 \neq \mathbf{x}_3$ , we can try restart GMRES)

the residuals $\mathbf{e} =$	ans $\mathbf{x}_3 =$
1.8227	-3.0710
-3.0710	1.3417
1.3417	0.2578
0.2578	0.1606
0.1606	2.2129
2.2129	2.2241
2.2241	1.8227

Restart after  $m=4$  iterations, and do not do any repetition.

the residuals $\mathbf{e} =$	ans $\mathbf{x}_4 =$
0.9539	-4.1038
-4.1038	1.0231
1.0231	0.1686
0.1686	0.3110
0.3110	1.8032
1.8032	1.8786
1.8786	1.9539

Restart after  $m=4$  iterations; repeat this 19 times; (Since  $\mathbf{x}_0 \neq \mathbf{x}_4$ , we can try restart GMRES)

the residuals $\mathbf{e} =$	answer $\mathbf{x}_4 =$
2.2210	-1.9378
-1.9378	1.0364
1.0364	0.0048
0.0048	0.0070
0.0070	0.7810
0.7810	2.0993
2.0993	3.2210

When the small size of  $H_m$  can not solve the problem, we may increase the size of  $H_m$  or restart GMRES.

Restart after  $m = 5$  iterations, repeat this 19 times;

the residuals $e=$	ans $x_5 =$
0.9069	-0.1641
-0.1641	0.1733
0.1733	-0.0467
-0.0467	0.1338
0.1338	0.0848
0.0848	0.1739
0.1739	1.9069

Restart after  $m = 5$  iterations, repeat this 99 times;

the residuals $e=$	ans $x_5 =$
0.0008	0.0011
0.0011	0.0010
0.0010	0.0005
0.0005	-0.0002
-0.0002	0.0011
0.0011	-0.0012
-0.0012	1.0008

Restart after  $m = 5$  iterations; repeat this 499 times;

the residuals $e=$	ans $x_5 = x =$
1.0e-018 *	0.0000
0	0.0000
0.5504	0.0000
0.8653	0.0000
-0.5548	0.0000
-0.1493	0.0000
-0.1962	1.0000
0.3928	

If we need to restart 499 times in order to get a good solution, it is hard to converge by using this initial guess. We may want to try other initial guesses (see data 2 (b)).

(b) random vector  $\in \mathbb{R}^7$ .

Restart after  $m=3$  iterations, repeat this 19 times;

$x_0 =$	the residuals $e =$	ans $x_3 =$
0.2190	-0.7050	0.6840
0.0470	0.6840	0.1474
0.6789	0.1474	0.4044
0.6793	0.4044	0.3582
0.9347	0.3582	0.4160
0.3835	0.4160	-0.0964
0.5194	-0.0964	0.2950

Restart after  $m=3$  iterations, repeat this 199 times;

$x_0 =$	the residuals $e =$	ans $x_3 =$
0.8310	-0.7531	0.4316
0.0346	0.4316	0.0616
0.0535	0.0616	0.4099
0.5297	0.4099	0.2601
0.6711	0.2601	0.6769
0.0077	0.6769	0.1240
0.3834	0.1240	0.2469

Restart after  $m=4$  iterations, repeat this 199 times;

$x_0 =$	the residual $e =$	ans $x_4 =$
0.0668	0.0094	0.0090
0.4175	0.0090	0.0017
0.6868	0.0017	-0.0057
0.5890	-0.0057	0.0090
0.9304	0.0090	-0.0053
0.8462	-0.0053	0.0022
0.5269	0.0022	1.0094

Restart after  $m=5$  iterations, repeat this 199 times;

$x_0 =$	residual $e =$	ans $x_5 =$
0.0920	0.0004	-0.0002
0.6539	-0.0002	0.0048
0.4160	0.0048	0.0005
0.7012	0.0005	0.0005
0.9103	0.0005	0.0027
0.7622	0.0027	-0.0013
0.2625	-0.0013	0.9996

Restart after  $m = 5$  iterations, repeat this 999 times;

$x_0 =$	residual $e =$	ans $x_5 =$
0.0475	1.0e-015 *	0.0000
0.7361	0.6661	0.0000
0.3282	0.2245	0.0000
0.6326	-0.0157	0.0000
0.7564	0.1134	0.0000
0.9910	-0.0726	0.0000
0.3653	0.0376	1.0000
	-0.1894	

Both data 2. (a) and (b) show that “restart GMRES method” converges slowly at iteration number 5. However, as long as we do not encounter a restart with the initial guess as the case of data 1, we should be able to get our exact solution.

$$8.3 \mathbf{H}_m(\cdot, m) = \sum_{i=1}^{m-1} a_i \mathbf{H}_m(\cdot, i). \text{ Then } r_{m-1}^g = r_m^g. \quad (11)$$

$$\text{If } \mathbf{H}_m(\cdot, m) = \sum_{i=1}^{m-1} a_i \mathbf{H}_m(\cdot, i), \text{ then } r_{m-1}^g = r_m^g.$$

*Proof:* Let  $\mathbf{h}_s = \mathbf{H}_m(\cdot, s)$ , then  $\mathbf{H}_m^g = \begin{bmatrix} \mathbf{h}_{m-1}^g & \mathbf{h}_m \\ 0 & \mathbf{h}_{m+1,m} \end{bmatrix}$ , where

$$\mathbf{H}_m = [\mathbf{h}_1 \ \mathbf{h}_2 \ \dots \ \mathbf{h}_{m-2} \ \mathbf{h}_{m-1} \ \mathbf{h}_m] \text{ and we can state (11) as } \mathbf{h}_m = \sum_{s=1}^{m-1} a_s \mathbf{h}_s.$$

$$\text{Let } \mathbf{h}_s^g = \mathbf{H}_m^g(\cdot, s), \text{ then } \mathbf{h}_s^g = \begin{bmatrix} \mathbf{h}_s \\ 0 \end{bmatrix} \text{ for } s = 1, 2, 3, \dots, m-1, \text{ and } \mathbf{h}_m^g = \begin{bmatrix} \mathbf{h}_m \\ \mathbf{h}_{m+1,m} \end{bmatrix}.$$

Since  $\mathbf{y}_m^g \in \mathbf{R}^m$  is the solution for  $\min\{\|\beta \mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m\|\}$ ,

$$\|\mathbf{r}_m^g\| = \|\beta \mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m\|$$

$$= \|\beta \mathbf{e}_{m+1} - [\mathbf{h}_1^g \ \mathbf{h}_2^g \ \dots \ \mathbf{h}_{m-1}^g \ \mathbf{h}_m^g] \begin{bmatrix} y_{m,1}^g \\ y_{m,2}^g \\ \vdots \\ \vdots \\ y_{m,m-1}^g \\ y_{m,m}^g \end{bmatrix}\|$$

$$= \|\beta \mathbf{e}_{m+1} - \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \dots & \mathbf{h}_{m-1} & \mathbf{h}_m \\ 0 & 0 & \dots & 0 & \mathbf{h}_{m+1,m} \end{bmatrix} \begin{bmatrix} y_{m,1}^g \\ y_{m,2}^g \\ \vdots \\ \vdots \\ y_{m,m-1}^g \\ y_{m,m}^g \end{bmatrix}\|$$

$$\begin{aligned}
&= \|\beta \mathbf{e}_{m+1} - \sum_{s=1}^{m-1} y_{m,s} \begin{bmatrix} \mathbf{h}_s \\ 0 \end{bmatrix} - y_{m,m} \begin{bmatrix} \mathbf{h}_m \\ \mathbf{h}_{m+1,m} \end{bmatrix}\| \\
&= \|\beta \mathbf{e}_{m+1} - \sum_{s=1}^{m-1} y_{m,s} \begin{bmatrix} \mathbf{h}_s \\ 0 \end{bmatrix} - y_{m,m} \left( \sum_{s=1}^{m-1} a_s \begin{bmatrix} \mathbf{h}_s \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{h}_{m+1,m} \end{bmatrix} \right)\| \\
&= \|\beta \mathbf{e}_{m+1} - \sum_{s=1}^{m-1} (y_{m,s} + y_{m,m} a_s) \begin{bmatrix} \mathbf{h}_s \\ 0 \end{bmatrix} + y_{m,m} \begin{bmatrix} 0 \\ \mathbf{h}_{m+1,m} \end{bmatrix}\|.
\end{aligned}$$

Since  $\left( \begin{bmatrix} \mathbf{h}_s \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{h}_{m+1,m} \end{bmatrix} \right) = 0$ , and since  $\mathbf{e}_{m+1} \perp \begin{bmatrix} 0 \\ \mathbf{h}_{m+1,m} \end{bmatrix}$ , if we let

$$\mathbf{w} = y_{m,m} \begin{bmatrix} 0 \\ \mathbf{h}_{m+1,m} \end{bmatrix} \text{ and let } \mathbf{v} = \beta \mathbf{e}_{m+1} - \sum_{s=1}^{m-1} (y_{m,s} + y_{m,m} a_s) \begin{bmatrix} \mathbf{h}_s \\ 0 \end{bmatrix}.$$

Then  $\|\mathbf{r}_m^g\| = \|\mathbf{v} + \mathbf{w}\|$ .

So,  $\|\mathbf{r}_m^g\|^2 = \|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 \geq \|\mathbf{v}\|^2 \geq \|\mathbf{r}_{m-1}^g\|^2$ .

Since  $\mathbf{v} = \beta \mathbf{e}_{m+1} - \sum_{s=1}^{m-1} (y_{m,s} + y_{m,m} a_s) \begin{bmatrix} \mathbf{h}_s \\ 0 \end{bmatrix}$  is in the class used to define  $\mathbf{r}_{m-1}$ ,

$$\|\mathbf{r}_m^g\| \geq \|\mathbf{r}_{m-1}^g\|.$$

But  $\|\mathbf{r}_{m-1}^g\| \geq \|\mathbf{r}_m^g\|$ , so  $\|\mathbf{r}_{m-1}^g\| = \|\mathbf{r}_m^g\|$ .

Hence, by (d) in section 1.8,  $\mathbf{r}_{m-1}^g = \mathbf{r}_m^g$ .

We use an example from [2] to show how it works out on this case.



**Example:** Let  $A$  be a skew-symmetric matrix of the form listed below.

Since  $A$  is skew symmetric, so is  $H_m$  for all  $m$ . Hence,  $H_m$  is singular for odd  $m$ , and the Arnoldi iterates  $x_m^A$  do not exist for  $m$  odd. For GMRES, we then have  $\|r_{2k}^g\| = \|r_{2k+1}^g\|$  for  $k = 1, 2, \dots, 24$ . These situations happen because the last column of  $H_{2k+1}$  is a linear combination of the previous columns. In the data we provide below:  $H_2$  and  $H_4$  are nonsingular, but  $H_3, H_5$  are singular because the columns are not linearly independent. So we should have  $x_2 = x_3$  and  $x_4 = x_5$ , according to the proof in 8.3.

$A =$	$b =$	$x_0 =$
0 -1 0 0 0 0 0	0.7071	0
1 0 -1 0 0 0 0	0	0
0 1 0 -1 0 0 0	0	0
0 0 1 0 -1 0 0	0	0
0 0 0 1 0 -1 0	0	0
0 0 0 0 1 0 -1	0	0
0 0 0 0 0 1 0	-0.7071	0

Restart after  $m = 2$  iterations, and do not do any repetition.

the residuals $r_2$	$H_2 =$ (non-singular)	ans $x_2 =$
=	0 -1	0
-0.3536	1 0	-0.3536
0		0
-0.3536		0
0		0
0.3536		-0.3536
0		0
0.3536		

Restart after  $m = 3$  iterations, and do not do any repetition.

the residuals $r_3$	$H_3 =$ (singular)	ans $x_3 =$
=	0 -1 0	0
-0.3536	1 0 -1	-0.3536
0	0 1 0	0
-0.3536		0
0		0
0.3536		-0.3536
0		0
0.3536		

We show that have  $x_2 = x_3$  in the former data and  $x_4 = x_5$  in the data we provide below:

Restart after  $m = 4$  iterations, and do not do any repetition.

$H_4$  (non-singular)=

0	-1.0000	0	0.0000
1.0000	0	-1.0000	0
0	1.0000	0	-1.4142
0	0	1.4142	0

the residuals  $r_4$ =

1.0e-015 \*  
-0.1110  
0  
0  
0  
0  
0  
0.1110

ans  $x_4 = x$  (the exact solution)=

0  
-0.7071  
0  
-0.7071  
0  
-0.7071  
0

Note: Since we happen to find exact solution from  $H_4$ , and since, according to the proof,

$r_4 = r_5$ ; we should have the exact solution from  $H_5$ . Hence,

Restart after  $m = 5$  iterations, and do not do any repetition.

$H_5$  (singular)=

0	-1.0000	0	0.0000	0
1.0000	0	-1.0000	0	1.0000
0	1.0000	0	-1.4142	0
0	0	1.4142	0	-1.4142
0	0	0	0.0000	0

the residuals  $r_5 =$

1.0e-015 \*  
-0.1110  
0  
0  
0  
0  
0  
0.1110

ans  $x_5 = x$ =

0  
-0.7071  
0  
-0.7071  
0  
-0.7071  
0

## Section 1.9 Restarting GMRES

To restart GMRES is to use the previous approximation  $\mathbf{x}_m^{\text{old}}$  as our new initial guess  $\mathbf{x}_m^{\text{new}}$  and obtain a new  $V_m$  (or new  $H_m$ ).

9.1 From the examples of Section 1.8, we showed that it was possible to have slow (or bad) convergence if we happened to choose a “bad initial guesses” which resulted in the singularities of  $H_m$ . In this situation, the best way to fix this problem is to find another initial guess. To restart GMRES is not to find an arbitrary new initial guess; but to use the previous approximation as the new initial guess since we can always find  $\mathbf{x}_m^g$  by applying GMRES, and since GMRES algorithm produces a sequence of non-increasing residuals (recall from Section 1.5, equation (10)).

*Proof:* Suppose we restart the GMRES after  $m$  iterations, and repeat this  $w-1$  times.

Let  $\mathbf{x}_{i,j}^g$  denote the approximate solution of the  $i^{\text{th}}$  time we restart the GMRES method after  $m$  iterations and then take  $j$  iterations. Let  $\mathbf{r}_{i,j}^g$  denote the corresponding residual.

Since to restart GMRES is to use the previous approximation as the new initial guess,  $\mathbf{x}_{0,m}^g = \mathbf{x}_{1,1}^g$ ,  $\mathbf{x}_{1,m}^g = \mathbf{x}_{2,1}^g$ ,  $\dots$ , and  $\mathbf{x}_{w-2,m}^g = \mathbf{x}_{w-1,1}^g$ .

Hence,  $\|\mathbf{r}_{0,m}^g\| = \|\mathbf{r}_{1,1}^g\|$ ,  $\|\mathbf{r}_{1,m}^g\| = \|\mathbf{r}_{2,1}^g\|$ ,  $\dots$ , and  $\|\mathbf{r}_{w-2,m}^g\| = \|\mathbf{r}_{w-1,1}^g\|$ .

So,  $\|\mathbf{r}_{0,1}^g\| \geq \|\mathbf{r}_{0,2}^g\| \geq \|\mathbf{r}_{0,3}^g\| \geq \dots \geq \|\mathbf{r}_{0,m}^g\| =$

$\|\mathbf{r}_{1,1}^g\| \geq \|\mathbf{r}_{1,2}^g\| \geq \|\mathbf{r}_{1,3}^g\| \geq \dots \geq \|\mathbf{r}_{1,m}^g\| =$

$\|\mathbf{r}_{2,1}^g\| \geq \|\mathbf{r}_{2,2}^g\| \geq \|\mathbf{r}_{2,3}^g\| \geq \dots \geq \|\mathbf{r}_{2,m}^g\| =$

$$\dots =$$

$$\|r_{w-1,1}^g\| \geq \|r_{w-1,2}^g\| \geq \|r_{w-1,3}^g\| \geq \dots \geq \|r_{w-1,m}^g\|.$$

9.2 However, there are always possibilities that  $\|r_{s-1,m}^g\| = \|r_{s,m}^g\|$ . This case may cause the restart method to fail to converge. If this situation happens, we can try a new initial guess, or even enlarge the size of  $m$ .

The following example shows that we have  $x_1^g \neq x$ ,  $x_2^g \neq x$ ,  $x_3^g \neq x$ , and  $x_4^g = x$  without restarting GMRES method, and we also can have  $x_3 = x$  with restarting GMRES method.

**Example:**

A =

1	-1	0	0	0	0	0
1	0	0	0	0	0	0
0	1	-1	0	0	0	0
0	0	0	1	0	0	0
0	0	0	0	1	0	0
0	0	0	0	0	1	0
0	0	0	0	0	0	1

b =

1
1
1
1
1
1
1

$x_0 =$

0
0
0
0
0
0
0

$r_0 =$

1
1
1
1
1
1
1

$r_0 A =$

0
1
0
1
1
1
1

$r_0 A^2 =$

-1
0
1
1
1
1
1

$r_0 A^3 =$

-1
-1
-1
1
1
1
1

Restart after  $m = 1$  iterations, and do not do any repetition.  
(we have  $x_1^g \neq x$  without restarting GMRES.)

the residuals e

= -1
0
-1
0
0
0
0

ans  $x_1$

= 1
1
1
1
1
1
1

$$H_1 = 0.8452$$

Restart after  $m = 2$  iterations, and do not do any repetition.

(we have  $\mathbf{x}_2^B \neq \mathbf{x}$  without restarting GMRES.)

the residuals $\mathbf{e}$	ans $\mathbf{x}_2$	
= -1	= 1	$H_2 =$ 0.8452 0.1604
0	1	0 1.4000
-1	1	
0	1	
0	1	
0	1	
0	1	

Restart after  $m = 3$  iterations, and do not do any repetition.

(we have  $\mathbf{x}_3^B \neq \mathbf{x}$  without restarting GMRES.)

the residuals $\mathbf{e} =$	ans $\mathbf{x}_3 =$	
-0.8000	1.5333	$H_3 =$
0.5333	1.3333	0.8452 0.1604 -0.0535
-0.2667	0.6000	0 1.4000 -0.2286
-0.1333	0.8667	0 0 1.2372
-0.1333	0.8667	
-0.1333	0.8667	
-0.1333	0.8667	

Restart after  $m = 4$  iterations, and do not do any repetition.

(we have  $\mathbf{x}_4^B = \mathbf{x}$  without restarting GMRES.)

the residuals $\mathbf{e} =$	ans $\mathbf{x}_4 = \mathbf{x} =$	
1.0e-015 *	1.0000	
0.8882	0.0000	
0.4441	-1.0000	$H_4 =$
-0.1110	1.0000	0.8452 0.1604 -0.0535 0.5071
0	1.0000	0 1.4000 -0.2286 -0.9939
0	1.0000	0 0 1.2372 -0.0522
0	1.0000	0 0 0 -0.6831

Next we show that with restarting GMRES method we can find the exact solution at an earlier iterate step.

Restart after  $m = 3$  iterations, repeat this 19 times; (Here we restart GMRES for 19 times)

the residuals $\mathbf{e} =$	ans $\mathbf{x}_3 = \mathbf{x} =$
1.0e-009 *	1.0000
-0.6868	0.0000
-0.5483	-1.0000
0.3156	1.0000
-0.0501	1.0000
-0.0501	1.0000
-0.0501	1.0000
-0.0501	1.0000

Restart after  $m = 3$  iterations, repeat this 49 times.

the residuals  $e =$

0  
0  
0  
0  
0  
0  
0  
0

ans  $x_j = x =$

1.0000  
0.0000  
-1.0000  
1.0000  
1.0000  
1.0000  
1.0000  
1.0000

## Section 1.10 Conclusions

10.1. There are two disadvantages of the Arnoldi algorithm, when we compare it to the GMRES algorithm.

- (a) One of the disadvantages of Arnoldi is that it does not produce a sequence of approximations whose residual norms are non-increasing, as GMRES does[2] (As we argued in Section 1.5, equation (10).) However, we can not find such a nice result for  $\mathbf{x}_m^{\wedge}$ .
- (b) A second perceived disadvantage of the Arnoldi algorithm is that it can break down and be unable to calculate an approximate solution. Partly for this reason, the GMRES method has been considered superior since it never breaks down in the way Arnoldi's algorithm does[2].

We can't solve  $\mathbf{x}_m^{\wedge} = \mathbf{x}_0 + \beta \mathbf{V}_m \mathbf{H}_m^{-1} \mathbf{e}_m$  for  $\mathbf{x}_m^{\wedge}$  in Arnoldi's when  $\mathbf{H}_m$  is singular. This can happen if either  $\mathbf{u}_{m+1} = \mathbf{0}$  so that  $\mathbf{v}_{m+1}$  can not be formed, or  $\mathbf{H}_{m_{\max}}$  is singular which means that the maximum number of Arnoldi steps has been taken, but the final iterate cannot be formed. The first situation has often been referred to as a happy break down, since  $\mathbf{u}_{m+1} = \mathbf{0}$  implies  $\mathbf{H}_m$  is nonsingular and  $\mathbf{x}_m^{\wedge}$  is the exact solution of  $\mathbf{Ax} = \mathbf{b}$  (on 8.1 of Section 1.8). The second case is more serious in that it causes a convergence failure of the algorithm. One possible recourse is to hope that  $\mathbf{H}_j$  is nonsingular for some  $j$  among  $1, 2, \dots, m_{\max} - 1$ . If such a  $j$  exists, we can compute  $\mathbf{x}_j^{\wedge}$  and then restart the algorithm using  $\mathbf{x}_j^{\wedge}$  as the new initial guess  $\mathbf{x}_0$ .

However, it may not always be possible to do this [2]. In this case ( examples on 8.2 and 8.3 of Section 1.8), even though the GMRES solution can always be calculated, the approximations are sometimes all the same until the very last one which gives the exact solution. In this situation, there is no use to restart GMRES; but we still can find our result at step  $n$ , by GMRES.

- (c) In appendix B, we show that  $\mathbf{r}_m^g \neq \mathbf{0}$  if and only if  $\mathbf{r}_m^A \neq \mathbf{0}$ . In other words, if we can not find the exact solution at the  $m^{\text{th}}$  iterate by Arnoldi method, GMRES method will not perform better at the same iteration step without restarting GMRES method.

## 10.2 What kinds of $A$ can we apply GMRES to solve $Ax = b$ ?

- (a) If we only require that the residual be small, then for any matrix  $A$  which is a non-singular square matrix can be used to find an approximation  $\mathbf{x}$  by applying GMRES.
- (b) Let norm of  $A$ ,  $\|A\|$ , be defined to be  $\sup\{|\langle A\mathbf{x}, \mathbf{x} \rangle| : \|\mathbf{x}\| = 1\}$ ,

$$\text{then } \|\mathbf{r}_m^g\| = \|\mathbf{b} - A\mathbf{x}_m^g\| = \|A\mathbf{x} - A\mathbf{x}_m^g\| = \|A(\mathbf{x} - \mathbf{x}_m^g)\| \leq \|A\| \|\mathbf{x} - \mathbf{x}_m^g\|.$$

$$\text{In other words, then } \|\mathbf{x} - \mathbf{x}_m^g\| \geq \|\mathbf{r}_m^g\| / \|A\|.$$

If we require  $\mathbf{x}$  to be very accurate, then we should be aware of the difference between

$$\|\mathbf{r}_m^g\| \text{ and } \|\mathbf{x} - \mathbf{x}_m^g\|. \text{ However, if } \|A\| \leq 1, \text{ then } \|\mathbf{x} - \mathbf{x}_m^g\| \geq \|\mathbf{r}_m^g\| / \|A\| \geq \|\mathbf{r}_m^g\|.$$

In other words, the smaller  $\|A\|$  is, the greater  $\|\mathbf{x} - \mathbf{x}_m^g\|$  may be. Hence, we should wisely choose the tolerance number i.e. decide acceptable values for the size of the residuals. If we set the tolerance number to be sufficiently small, any square,



nonsingular matrix will be able to find a approximate  $\mathbf{x}$  by applying GMRES.

(c) In the end of this chapter, we provide a example to show how GMRES method handles a bigger matrix like a  $50 \times 50$  Hilbert matrix. Matrices that are sufficiently large cannot be handled with conventional techniques like Gaussian Elimination or using inverse  $\mathbf{A}$  etc.....

The example is shown below:

**Example:** Let  $\mathbf{A}$  be a  $50 \times 50$  Hilbert matrix, where  $a_{ij} = 1/(i+j-1)$ .

$$\mathbf{b} = \mathbf{A} \mathbf{x} \quad \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \\ 1 \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \\ 1 \end{bmatrix} \quad \text{is a vector of length 50, and}$$

$$\mathbf{x}_0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad \text{is a vector of length 50;}$$

then, we can expect that the exact solution to be

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \\ 1 \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \\ 1 \end{bmatrix} \quad \text{is a vector of length 50.}$$

In this example, we try restart after  $m = 7$  iterations and repeat this 20, 50, 100, and 200 times; restart after  $m = 9$  iterations and repeat this 20 and 50 times; and restart after  $m = 10$  iterations and repeat this 20 times. The table below shows how the GMRES works between different iteration numbers and different times of restart:

**Table for the residuals  $r_m^g$  :**

The range of The norms of The residuals $r_m^g$	restart 20 times	Restart 50 times	restart 100 times	Restart 200 times
Iteration number $m = 7$	$1.0 \times 10^{-10}$	$1.0 \times 10^{-11}$	$1.0 \times 10^{-11}$	$1.0 \times 10^{-11}$
Iteration number $m = 9$	$1.0 \times 10^{-14}$	$1.0 \times 10^{-14}$		
Iteration number $m = 10$	$1.0 \times 10^{-15}$			

**The residuals tell how close the approximate solutions are to the exact solution.**

**(d) Next, we want to apply the GMRES method to solve the wavelet transform coefficient vector. In chapter 2, we provide some common knowledge of the wavelets, details about Haar wavelet; and then show how the GMRES method can be applied to wavelets.**

## **CHAPTER 2. WAVELETS**

### **Section 2.1 An introduction to wavelets**

(This section is based on **An introduction to Wavelets**, by Amara Graps [14])

Wavelets are functions that satisfy certain mathematical requirements and are used in representing data or other functions. This idea is not new. Approximation using superposition of functions has existed since the early 1800's, when Joseph Fourier discovered that he could superpose sines and cosines to represent other functions.

He asserted that any  $2\pi$ -periodic function  $f(x)$  is the sum  $a_0 + \sum_{k=1}^m (a_k \cos(kx) + b_k \sin(kx))$

of its Fourier series. The coefficients  $a_0$ ,  $a_k$ , and  $b_k$  are calculated by

$$a_0 = (1/2\pi) \int_0^{2\pi} f(x) dx, \quad a_k = (1/\pi) \int_0^{2\pi} f(x) \cos(kx) dx, \quad \text{and} \quad b_k = (1/\pi) \int_0^{2\pi} f(x) \sin(kx) dx.$$

However, in wavelet analysis, the scale that we use to look at data plays a special role.

Wavelet algorithms process data at different scales or resolutions. If we look at a signal with a large window, we would notice gross features. Similarly, if we look at a signal with a small window, we would notice small features. The result in wavelet analysis is to see both the forest and the trees, so to speak.

This makes wavelets interesting and useful. For many decades, scientists have wanted more appropriate functions than the sines and cosines which comprise the bases of Fourier analysis, to approximate choppy signals. By their definition, these functions

are non-local (and stretch out to infinite). They therefore do a very poor job in approximating sharp spikes. But with wavelet analysis, we can use approximating functions that are contained neatly in finite domains. Wavelets are well-suited for approximating data with sharp discontinuities.

Understanding the concepts of basis functions and scale-varying basis functions is key to understanding wavelets.[14]

### 1.1 What are Basis Functions?

It is simpler to explain a basis function if we move out of the realm of analog (functions) and into the realm of digital (vectors)[15]. Every two-dimensional vector  $(x,y)$  is a combination of the vector  $(1,0)$  and  $(0,1)$ . These two vectors are the basis vectors for  $(x,y)$ . Why? Notice that  $x$  multiplied by  $(1,0)$  is the vector  $(x,0)$ , and  $y$  multiplied by  $(0,1)$  is the vector  $(0,y)$ . The sum is  $(x,y)$ .

The best basis vectors have the valuable extra property that the vectors are perpendicular, or orthogonal to each other. For the basis  $(1,0)$  and  $(0,1)$ , this criteria is satisfied.

Now let's go back to the analog world, and see how to relate these concepts to basis functions. Instead of the vector  $(x,y)$ , we have a function  $f(x)$ . Imagine that  $f(x)$  is a musical tone, say the note A in a particular octave. We can construct A by adding sines and cosines using combinations of amplitudes and frequencies. The sines and cosines are the basis functions in this example, and the elements of Fourier synthesis. For the sines

and cosines chosen, we can set the additional requirement that they be orthogonal. How? By choosing the appropriate combination of sine and cosine function terms whose inner product adds up to zero. The particular set of functions that are orthogonal and that construct  $f(x)$  are our orthogonal basis functions for this problem.

## 1.2 What are Scale-Varying Basis Functions?

A basis function varies in scale by chopping up the same function or data space using different scale sizes. For example, imagine we have a signal over the domain from 0 to 1. We can divide the signal with two step functions that range from 0 to  $\frac{1}{2}$  and  $\frac{1}{2}$  to 1. Then we can divide the original signal again using four step functions from 0 to  $\frac{1}{4}$ ,  $\frac{1}{4}$  to  $\frac{1}{2}$ ,  $\frac{1}{2}$  to  $\frac{3}{4}$ , and  $\frac{3}{4}$  to 1. And so on. Each set of representations code the original signal with a particular resolution or scale.

In this chapter, we use Haar basis to construct a function.

## Section 2.2 Haar basis

In this section we will show how an arbitrary function  $f$  can be approximated by a linear combination of Haar wavelets. Below, we use the information from [3] to describe Haar basis and Haar wavelets.

**Lemma 1.**  $\Psi_{m,n}(x) = 2^{m/2}\Psi(2^m x - n)$  constitute an orthonormal basis for  $L^2(\mathcal{R})$ , where

$$\Psi(x) = \begin{cases} 1 & 0 \leq x < 1/2 \\ -1 & 1/2 \leq x < 1 \\ 0 & \text{otherwise} . \end{cases}$$

*Proof:* Since support  $(\Psi_{m,n}) = [2^{-m}n, 2^{-m}(n+1)]$ , it follows that two Haar wavelets of the same scale (same value of  $m$ ) never overlap, so that  $\langle \Psi_{m,n}, \Psi_{m,n'} \rangle = \delta_{n,n'}$ .

Overlapping supports are possible if the two wavelets have different sizes, as in (figure 2.2.1.) It is easy to check, however, that if  $m < m'$ , then support  $(\Psi_{m,n})$  lies wholly within a region where  $\Psi_{m',n'}$  is constant (as on the figure). It follows that the inner product of  $\Psi_{m,n}$  and  $\Psi_{m',n'}$  is then proportional to the integral of  $\Psi$  itself, which is zero.

Any  $f$  in  $L^2(\mathcal{R})$  can be arbitrarily well approximated by a function with compact support which is piecewise constant on the  $[2^{-m}n, 2^{-m}(n+1)[$  (it suffices to take the support and  $m$  large enough). We can therefore restrict ourselves to such piecewise constant functions only.

We describe here a recursive method for representing  $f$ :

Assume  $f$  to be supported on  $[-2^{m_1}, 2^{m_1}]$ , and to be piecewise constant on the intervals  $[2^{-m_0}n, 2^{-m_0}(n+1)[$ , where  $m_1$  and  $m_0$  can both be arbitrary large (figure 2.2.2). Let us denote the constant value of  $f^0 = f$  on  $[2^{-m_0}n, 2^{-m_0}(n+1)[$  by  $f_n^0$ . We now represent  $f^0$  as a sum of  $f^0 = f^1 + \delta^1$ , where  $f^1$  is an approximation to  $f^0$  which is piecewise constant over intervals twice as large as originally, i.e.  $f^1 | [2^{-m_0+1}n, 2^{-m_0+1}(n+1)[ = \text{constant} = f_n^1$ . The values  $f_n^1$  are given by the averages of the two corresponding constant values for  $f^0$ ,  $f_n^1 = \frac{1}{2} (f_{2n}^0 + f_{2n+1}^0)$  (see figure 2.2.3). The function  $\delta^1$  is piecewise constant with the same stepwidth as  $f^0$ ; one immediately has (as shown in figure 2.2.3.)

$$\delta_{2n}^1 = f_{2n}^0 - f_n^1 = \frac{1}{2} (f_{2n}^0 + f_{2n+1}^0)$$

and

$$\delta_{2n+1}^1 = f_{2n+1}^0 - f_n^1 = \frac{1}{2} (f_{2n+1}^0 + f_{2n}^0) = -\delta_{2n}^1.$$

It follows that  $\delta^1$  is a linear combination of scaled and translated Haar functions:

$$\delta^1 = \sum_{n=s}^t \delta_{2n}^1 \Psi(2^{m_0-1}x - n), \text{ where } s = -2^{m_1+m_0-1} + 1 \text{ and } t = 2^{m_1+m_0-1}.$$

We have therefore written  $f$  as  $f = f^0 = f^1 + \sum_n C_{-m_0+1,n} \Psi_{-m_0+1,n}$ , where  $f^1$  is of the same type as  $f^0$ , but with stepwidth twice as large. (see figure 2.2.3.)

Note: (1)  $f_n^1 = \frac{1}{2} (f_{2n}^0 + f_{2n+1}^0)$  is the average of the sum, can be denoted by "s".

$$(2) \quad \delta_{2n}^1 = f_{2n}^0 - f_n^1 = \frac{1}{2} (f_{2n}^0 + f_{2n+1}^0)$$

and

$$\delta_{2n+1}^1 = f_{2n+1}^0 - f_n^1 = \frac{1}{2} (f_{2n+1}^0 + f_{2n}^0) = -\delta_{2n}^1$$



are the difference, can be denoted by “d”. When we average the sum, we keep the difference as the detail; so “d” also denotes “detail”. [15]

(3) The coefficients  $C_{-m0+1,n}$  of  $\sum_n C_{-m0+1,n} \Psi_{-m0+1,n}$  as in the equation

$$f = f^0 = f^1 + \sum_n C_{-m0+1,n} \Psi_{-m0+1,n}$$

are called the Haar wavelet coefficients.

Note that the coefficients also represent the differences.

(4) If the differences (the wavelet coefficients) are too small (smaller than a threshold number), we normally set these differences (the wavelet coefficients) towards zero. We call this procedure “de-noising”. We smooth the data by de-noising the data. [16]

We can apply the same trick to  $f^1$ , so that  $f^1 = f^2 + \sum_n C_{-m0+2,n} \Psi_{-m0+2,n}$ , with  $f^2$  still supported on  $[-2^{m1}, 2^{m1}]$ , but piecewise constant on the even larger intervals  $[2^{-m0+2}n, 2^{-m0+2}(n+1)[$ . We can keep going like this, until we have

$$f = f^{m0+m1} + \sum_{k=-m0+1}^{m1} \sum_n C_{k,n} \Psi_{k,n}$$

Here  $f^{m0+m1}$  consists of two constant pieces (see figure 2.2.3), with

$f^{m0+m1} | [0, 2^{m1}[ = f_0^{m0+m1}$  is equal to the average of  $f$  over  $[0, 2^{m1}[$ , and

$f^{m0+m1} | [-2^{m1}, 0[ = f_{-1}^{m0+m1}$  is equal to the average of  $f$  over  $[-2^{m1}, 0[$ .

Even though we have “filled out” the whole support of  $f$ , we can still keep going with our “average” trick: nothing stops us from widening our horizon from  $2^{m1}$  to  $2^{m1+1}$ , and writing  $f^{m1+m2} = f^{m1+m2+1} + \delta^{m1+m2+1}$ , where

$$f^{m1+m2+1} | [0, 2^{m1+1}[ = \frac{1}{2} f_0^{m1+m2},$$

$$f^{m_1+m_2+1} | [-2^{m_1+1}, 0[ = \frac{1}{2} f_{\cdot 1}^{m_1+m_2} \quad \text{and}$$

$$\delta^{m_1+m_2} = \frac{1}{2} f_0^{m_1+m_2} \Psi(2^{-m_1-1} x) - \frac{1}{2} f_{\cdot 1}^{m_1+m_2} \Psi(2^{-m_1-1} x + 1)$$

(see figure 2.2.3). This can again be repeated, leading to

$$f = f^{m_0+m_1+h} + \sum_{k=-m_0+1}^{m_1+h} \sum_n C_{k,n} \Psi_{k,n},$$

where support  $(f^{m_0+m_1+h}) = [-2^{m_1+h}, 2^{m_1+h}]$ , and

$$f^{m_0+m_1+h} | [0, 2^{m_1+h}[ = 2^{-h} f_0^{m_0+m_1},$$

$$f^{m_0+m_1+h} | [-2^{m_1+h}, 0[ = 2^{-h} f_{\cdot 1}^{m_0+m_1}.$$

It follows immediately that

$$\begin{aligned} & (\text{norm}(f - \sum_{k=-m_0+1}^{m_1+h} \sum_n C_{k,n} \Psi_{k,n}))^2 \\ &= (\text{norm}(f^{m_0+m_1+h}))^2 \\ &= 2^{-h/2} 2^{m_1/2} [(\text{norm}(f_0^{m_0+m_1}))^2 + (\text{norm}(f_{\cdot 1}^{m_0+m_1}))^2]^{1/2}, \end{aligned}$$

which can be made small by taking sufficiently large  $K$ . If we continue until our support intervals are the same as the intervals where  $f$  is a constant function, we express  $f$  in as a linear sum of the wavelets exactly. [3]

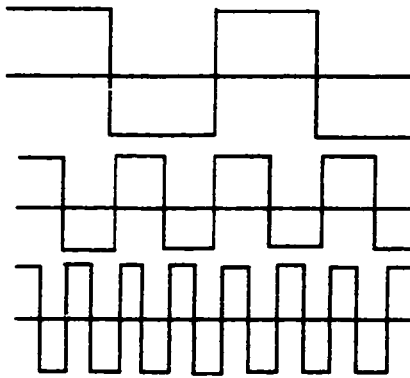


figure 2.2.1

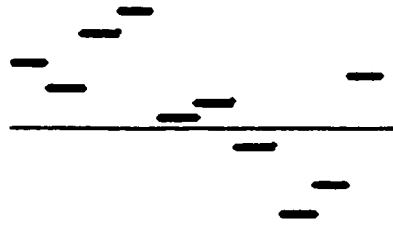


figure 2.2.2

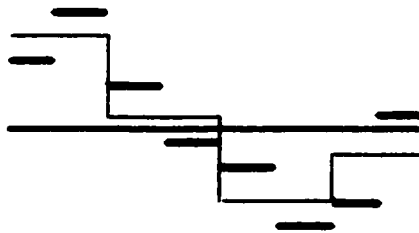


figure 2.2.3

### Section 2.3 Wavelet transforms for discrete data

Because the original signal or function can be represented in terms of a wavelet expansion (using coefficients in a linear combination of the wavelet functions), data operations can be performed using just the corresponding wavelet coefficients. And if you further choose the best wavelets adapted to your data, or truncate the coefficients below a threshold, your data is sparsely represented. This sparse coding makes wavelets an excellent tool in the field of data compression. [14]

The argument in Section 2.2 has implicitly used a “multiresolution” approach: We have written successive coarser and coarser approximations to  $f$  (the  $f^m$ , averaging  $f$  over larger and larger intervals), and at every step we have written the difference between the approximation with resolution  $2^{m-1}$ , and the next coarser level, with resolution  $2^m$ , as a linear combination of the  $\Psi_{m,n}$ . [8] We identify

$$\Psi_{m,n}(x) = 2^{m/2} \Psi(2^m x - n) \text{ to be}$$

$$\Psi_{m,n}(x) = 2^{m/2} (0, 0, 0, \dots, 1, 1, \dots, 1, -1, -1, \dots, -1, 0, 0, \dots, 0),$$

one of the column vectors of the matrix  $A$  ( see table 1 and table 2).

Given a set of  $2^k$  discrete data points, which is supported on  $[0, 2^k]$  with uniform grids. We can construct  $f$  by letting  $f_i$  representing these values on each unit interval of interval  $[0, 2^k]$ . Then  $f^0$  is  $f$ , and  $f^k = 2^{-k/2} \sum_{i=1}^k f_i$ , a constant function, is the coarsest approximation to  $f$ . Let  $r = 1, 2, \dots, k$  and  $n = 1, 2, \dots, 2^{k-r}$ , we can also write  $f$  as

$$\begin{aligned}
\mathbf{f} &= \mathbf{f}^0 \\
&= \mathbf{f}^1 + \sum_{n=1}^a C_{1,n} \Psi_{1,n}, && \text{(where } a = 2^{k-1} \text{)} \\
&= \mathbf{f}^2 + \sum_{n=1}^b C_{2,n} \Psi_{2,n} + \sum_{n=1}^a C_{1,n} \Psi_{1,n} && \text{(where } b = 2^{k-2} \text{)} \\
&= \mathbf{f}^3 + \sum_{n=1}^c C_{3,n} \Psi_{3,n} + \sum_{n=1}^b C_{2,n} \Psi_{2,n} + \sum_{n=1}^a C_{1,n} \Psi_{1,n} && \text{(where } c = 2^{k-3} \text{)} \\
&= \dots \\
&= \mathbf{f}^r + \sum_{n=1}^s C_{r,n} \Psi_{r,n} + \dots + \sum_{n=1}^b C_{2,n} \Psi_{2,n} + \sum_{n=1}^a C_{1,n} \Psi_{1,n} && \text{(where } s = 2^{k-r} \text{)} \\
&= \dots \\
&= \mathbf{f}^k + \sum_{n=1}^l C_{k,n} \Psi_{k,n} + \dots + \sum_{n=1}^s C_{r,n} \Psi_{r,n} + \dots + \sum_{n=1}^b C_{2,n} \Psi_{2,n} + \sum_{n=1}^a C_{1,n} \Psi_{1,n} \\
&= \mathbf{f}^k + \\
&\quad C_{k,1} \Psi_{k,1} + \\
&\quad C_{k-1,1} \Psi_{k-1,1} + C_{k-1,2} \Psi_{k-1,2} + \\
&\quad C_{k-2,1} \Psi_{k-2,1} + C_{k-2,2} \Psi_{k-2,2} + C_{k-2,3} \Psi_{k-2,3} + C_{k-2,4} \Psi_{k-2,4} + \\
&\quad \dots + \\
&\quad C_{r,1} \Psi_{r,1} + C_{r,2} \Psi_{r,2} + \dots + C_{r,s} \Psi_{r,s} + && \text{(where } s = 2^{k-r} \text{)} \\
&\quad \dots + \\
&\quad C_{2,1} \Psi_{2,1} + C_{2,2} \Psi_{2,2} + \dots + C_{2,b} \Psi_{2,b} + && \text{(where } b = 2^{k-2} \text{)} \\
&\quad C_{1,1} \Psi_{1,1} + C_{1,2} \Psi_{1,2} + \dots + C_{1,a} \Psi_{1,a} . && \text{(where } a = 2^{k-1} \text{)}
\end{aligned}$$

Recall we can associate each  $\Psi_{m,n}$  with a vector of form:

$(0, 0, \dots, 0, 1, 1, \dots, 1, 1, -1, -1, \dots, -1, -1, 0, 0, \dots, 0, 0)$ . Likewise, the original  $f$  can be represented as a  $2^k$  vector,  $\mathbf{f}$ . Hence this is a linear system. Since there are totally

$1 + \sum_{r=0}^{k-1} 2^r = 1 + (2^k - 1) = 2^k$  items; for  $n = 2, 3, \dots, 2^k$ , we can rename all the Haar

wavelets, the associated vector  $\Psi_{i,j}$ , as a vector  $\mathbf{a}_n$ , and all the corresponding coefficients as scalar  $d_n$ .

In addition, let  $\mathbf{f}^k = d_1 \mathbf{a}_1$  where  $\mathbf{a}_1 = [1, 1, 1, \dots, 1] / \|[1, 1, 1, \dots, 1]\|$ . Then we have a matrix  $\mathbf{A}$  such that the  $n^{\text{th}}$  column of  $\mathbf{A}$  is vector  $\mathbf{a}_n$ , and a vector  $\mathbf{d}$  with entry  $d_n$ .

Since  $\mathbf{f}$  is identified with the  $2^k$  vector with coordinate  $f_j$  using vector notation, then  $\mathbf{f} = \mathbf{A}\mathbf{d}$ . Now, we wish to find  $\mathbf{d}$ , or, most likely, an approximation to  $\mathbf{d}$ . To find an approximation to  $\mathbf{d}$ , we can apply GMRES to approximate  $\mathbf{d}$ .

Since  $\mathbf{A}$  is orthonormal, we can also find  $\mathbf{d}$  by  $\mathbf{d} = \mathbf{A}^T \mathbf{f}$ . However we will focus on using the GMRES method to illustrate how the GMRES works in this application. In this paper, we provide an example to show how the GMRES method helps transform a data vector  $\mathbf{f}$  to be a wavelet transform coefficient vector,  $\mathbf{d}$ , by a  $128 \times 128$  ( $128 = 2^7$ ) matrix,  $\mathbf{A}$ . When we apply the GMRES method, we can choose some  $m < 128$ , so

$$\mathbf{x}_m^g = \mathbf{x}_0 + \mathbf{V}_m \mathbf{y}_m, \text{ where } \mathbf{y}_m^g \in \mathbf{R}^m \text{ (equation (7)) such that}$$

$$\|\mathbf{r}_m^g\| = \min\{\|\beta \mathbf{e}_{m+1} - \mathbf{H}_m^g \mathbf{y}_m\|\}, \text{ for } \mathbf{y}_m \in \mathbf{R}^m \text{ (equation 9)).}$$

In other words, using the GMRES method, we only have to deal with a small matrix instead of a  $128 \times 128$  matrix. We hope the residual  $\|\mathbf{r}_m^g\|$  is very small; let's see how

the example works.

Before we show the example, we first list 3 different sizes of A's (table 1) and the program (table 2) which constructs A.

**Table 1:**

$A_4 =$

0.5000	-0.5000	-0.7071	0
0.5000	-0.5000	0.7071	0
0.5000	0.5000	0	-0.7071
0.5000	0.5000	0	0.7071

$A_8 =$

0.3536	-0.3536	-0.5000	0	-0.7071	0	0	0
0.3536	-0.3536	-0.5000	0	0.7071	0	0	0
0.3536	-0.3536	0.5000	0	0	-0.7071	0	0
0.3536	-0.3536	0.5000	0	0	0.7071	0	0
0.3536	0.3536	0	-0.5000	0	0	-0.7071	0
0.3536	0.3536	0	-0.5000	0	0	0.7071	0
0.3536	0.3536	0	0.5000	0	0	0	-0.7071
0.3536	0.3536	0	0.5000	0	0	0	0.7071

$A_{16} =$

0.2500	-0.2500	-0.3536	0	-0.5000	0	0	0	-0.7071	0	0	0	0	0	0	0
0.2500	-0.2500	-0.3536	0	-0.5000	0	0	0	0.7071	0	0	0	0	0	0	0
0.2500	-0.2500	-0.3536	0	0.5000	0	0	0	0	-0.7071	0	0	0	0	0	0
0.2500	-0.2500	-0.3536	0	0.5000	0	0	0	0	0.7071	0	0	0	0	0	0
0.2500	-0.2500	0.3536	0	0	-0.5000	0	0	0	0	-0.7071	0	0	0	0	0
0.2500	-0.2500	0.3536	0	0	-0.5000	0	0	0	0	0.7071	0	0	0	0	0
0.2500	-0.2500	0.3536	0	0	0.5000	0	0	0	0	0	-0.7071	0	0	0	0
0.2500	-0.2500	0.3536	0	0	0.5000	0	0	0	0	0	0.7071	0	0	0	0
0.2500	0.2500	0	-0.3536	0	0	-0.5000	0	0	0	0	0	-0.7071	0	0	0
0.2500	0.2500	0	-0.3536	0	0	-0.5000	0	0	0	0	0	0.7071	0	0	0
0.2500	0.2500	0	-0.3536	0	0	0.5000	0	0	0	0	0	0	-0.7071	0	0
0.2500	0.2500	0	-0.3536	0	0	0.5000	0	0	0	0	0	0	0.7071	0	0
0.2500	0.2500	0	0.3536	0	0	0	-0.5000	0	0	0	0	0	0	-0.7071	0
0.2500	0.2500	0	0.3536	0	0	0	-0.5000	0	0	0	0	0	0	0.7071	0
0.2500	0.2500	0	0.3536	0	0	0	0.5000	0	0	0	0	0	0	0	-0.7071
0.2500	0.2500	0	0.3536	0	0	0	0.5000	0	0	0	0	0	0	0	0.7071

**Table 2:**

The following is a program to construct the matrix, **A**, by matlab.

```

for m = 1 : 1 : 2k,
  for n = 1 : 1 : 2k,
    if n == 1;
      A(m,1) = 1;
    else
      if log2(n) == fix(log2(n)),
        s = log2(n) - 1;
      else
        s = fix(log2(n));
      end
      y = m × (2s-k) - (n - 2s) + 1;
      if (0 < y) & (y <= 0.5),
        A(m,n) = -2(s-k)/2;
      elseif (0.5 < y) & (y <= 1.0),
        A(m,n) = 2(s-k)/2;
      else
        A(m,n) = 0;
      end
    end
  end
end
end

```

The following example shows that the matrix **A** transforms the original data to the coarsest approximation in one step; so that we can easily de-noise the data (by setting the small coefficients to be zero). This example also shows that the GMRES method helps to find a wavelet transform coefficient vector with a error as small as  $1.3323 \times 10^{-15}$ .

**Example:** Let **f** be a  $128 \times 1$  data vector (see table 3: Original data) such that

$$\mathbf{f}(n) = 1.1 + 0.01 \times \sin(100000 \times n), \text{ where } n = 1, 2, \dots, 32;$$

$$\mathbf{f}(n) = 1.5 + 0.01 \times \sin(100000 \times n), \text{ where } n = 31, 32, \dots, 64; \text{ and}$$

$$\mathbf{f}(n) = 2.1 + 0.01 \times \sin(100000 \times n), \text{ where } n = 65, 66, \dots, 128.$$



**First step:** Construct a  $128 \times 128$  matrix  $A$ .

We wish to approximate  $x$  so that  $f = Ax$  ( note that this  $x$  is exactly the  $d$  that we mentioned earlier in this section.)

**Second step:** We restart the GMRES after  $m$  iterations and repeat this a few times. Call the final approximation  $x_m^g$  and the corresponding residual  $r_m^g$ . We did a few experiments and list all the residuals from each experiment below:

**Norm of final residuals  $r_m^g$  :**

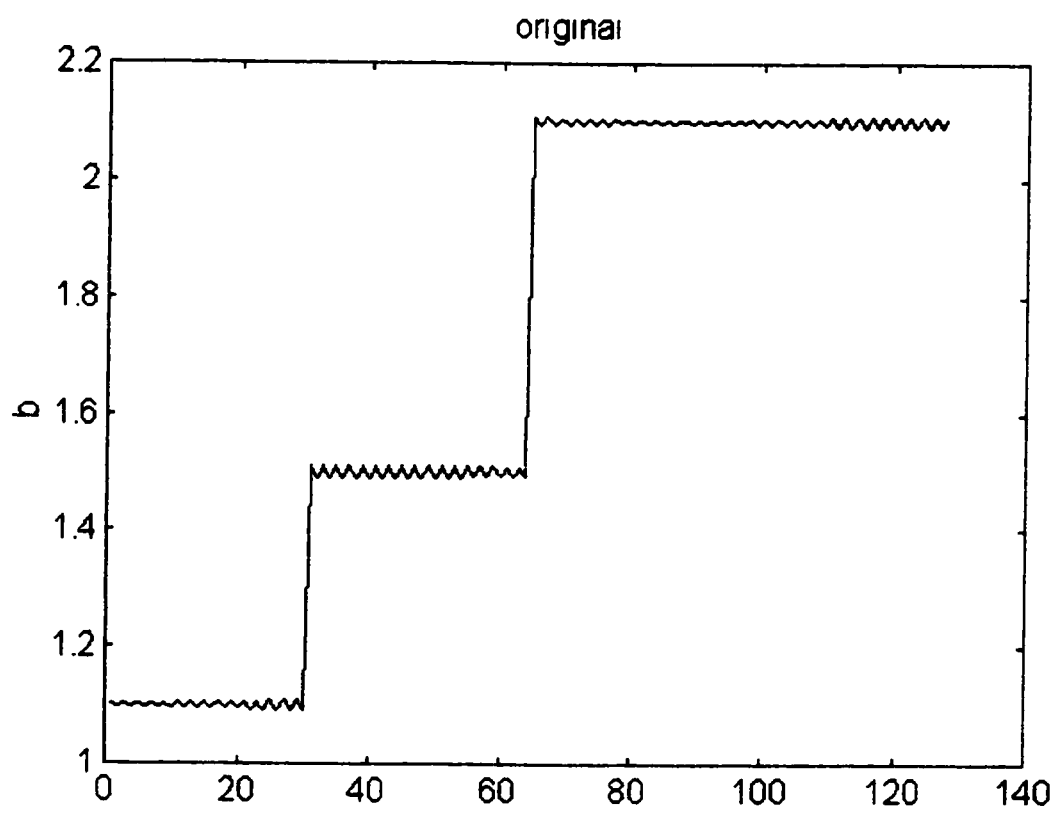
repetitions $m \setminus$	repeat 0 time	repeat 1 time	repeat 5 Times	repeat 10 times	repeat 20 Times
$m = 10$	4.6690	2.2586	1.6637	1.4840	0.7351
$m = 20$	1.7563	0.2689	$8.8594 \times 10^{-4}$	$2.8764 \times 10^{-6}$	$3.4309 \times 10^{-11}$
$m = 25$	1.3796	0.1095	$5.9927 \times 10^{-6}$	$3.1029 \times 10^{-11}$	$1.3867 \times 10^{-15}$
$m = 30$	0.0489	$1.2811 \times 10^{-4}$	$1.1835 \times 10^{-14}$	$1.6012 \times 10^{-15}$	
$m = 40$	$3.7962 \times 10^{-12}$	$1.3323 \times 10^{-15}$			
$m = 50$	$1.3459 \times 10^{-14}$	$1.4895 \times 10^{-15}$			
$m = 64$	$1.5715 \times 10^{-14}$	$1.3688 \times 10^{-15}$			

We chose the most accurate  $x_m^g$  to be our  $d$  (table 4), the wavelet transform coefficient vector.

**Third step:** we set a threshold to smooth the coefficients (de-noise)[11] and have a new  $d$ , denoted by  $d_d$  (table 5). In this example we set the entries of  $d$  to be zero if they are smaller than 0.1 .

Recall from section 1: note (1) and (2), the entries of  $d$  are all the differences except the first entry. In other words, we only keep the average of the sum, if the difference is very small.

**Fourth step:** we find our new data  $f_d$  (table 6) with no noise by  $f_d = Ad_d$ .



**table 3:** original data vector **f** and picture.

**d =**

19.3045	0.0000	0.0000	-0.0123		
4.4559	0.0000	0.0000	-0.0128	-0.0032	
1.5007	0.0000	0.0000	-0.0132	-0.0022	
-0.0002	0.0000	0.0000	-0.0135	-0.0012	
0.1416	0.0000	0.0000	-0.0138	-0.0002	
0.0003	0.0000	0.0000	-0.0140	0.0008	
00.0001	0.0000	0.0000	-0.0141	0.0018	
-0.0002	0.0000	0.0000	-0.0141	0.0028	
0.0000	0.0000	0.0000	-0.0141	0.0038	
0.2001	0.0000	0.0000	-0.0140	0.0048	
0.0001	0.0000	0.0000	-0.0139	0.0057	
0.0001	0.0000	0.0000	-0.0136	0.0066	
0.0001	0.0000	-0.0008	-0.0133	0.0075	
0.0000	0.40000	-0.0018	-0.0129	0.0083	
-0.0001	0.0000	-0.0028	-0.0125	0.0091	
-0.0001	0.0000	-0.0037	-0.0120	0.0099	
0.0000	0.0000	-0.0047	-0.0114	0.0106	
0.0000	0.0000	-0.0057	-0.0108	0.0112	
0.0000	0.0000	-0.0066	-0.0101	0.0118	
0.2829	0.0000	-0.0074	-0.0094	0.0123	
0.0000	continue	-0.0083	-0.0086	continue	0.0128
0.0000	0.0000	-0.0091	-0.0078	0.0132	
0.0000	0.0000	-0.0098	-0.0069	0.0135	
0.0000	0.0000	-0.0105	-0.0060	0.0138	
0.0000	0.0000	-0.0112	-0.0051	0.0140	
0.0000	0.0000	-0.0118	-0.0042		

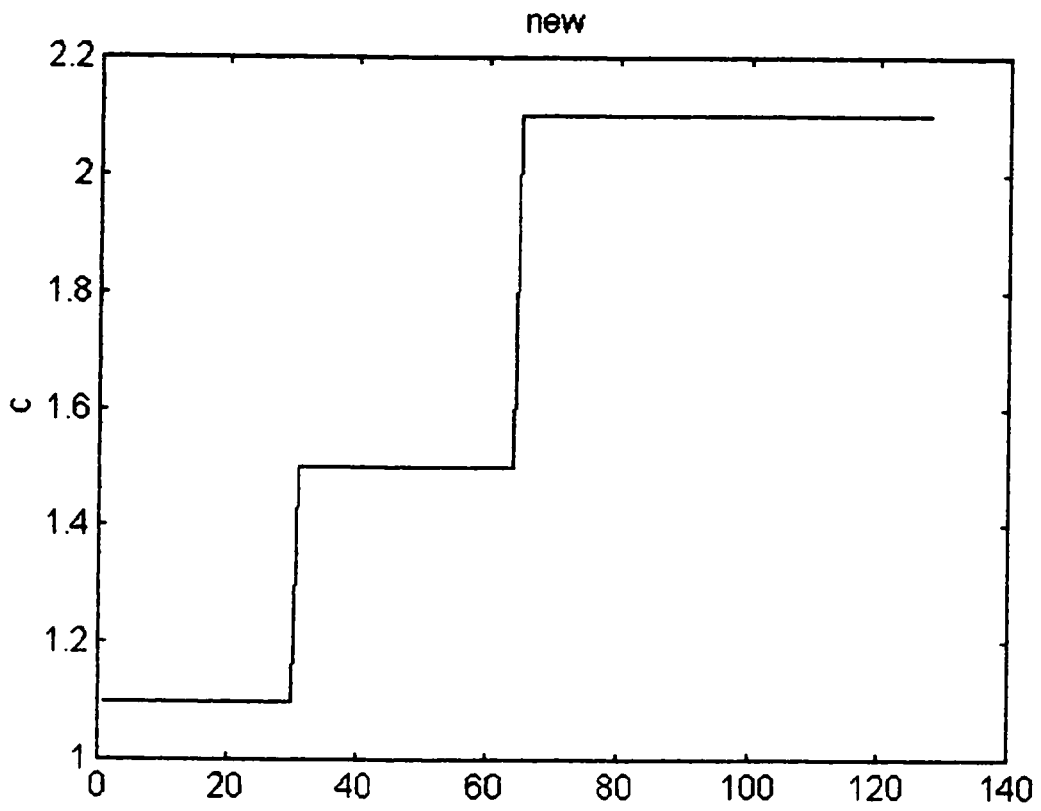
**table 4:**  
The wavelet transform coefficient vector, **d**, before de-noise.

$d_d =$

19.3045	0	0	0	0
4.4559	0	0	0	0
1.5007	0	0	0	0
0	0	0	0	0
0.1416	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0.2001	0	0	0	0
0	0	0	0.00000	0
0	0	0	0	0
0	0	0.00000	0	0.00000
0	0.40000	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0.2829	0	0	0	0
0	continue	0	continue	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

Note: 0.00000 shown here are exactly equal to zero.

**table 5:**  
The new wavelet transform coefficients,  $d_d$ , after de-noise.



**table 6:**  
The new data,  $f_d$ , and the picture after "de-noise".

## **Section 2.4 Conclusion and further studies**

GMRES method is based on a classical scheme due to Arnoldi (1951) that constructs an orthonormal basis of a space called the Krylov subspace. The GMRES (Generalized Minimal Residual) method is designed to minimize the norm of the residual vector  $\mathbf{b}-\mathbf{Ax}$  over all vectors of the affine Krylov subspace.

Restarting the GMRES method will produce a sequence of non-increasing residuals. However, in the present paper, we have shown, by geometry,

- (a) that some initial guesses will make the GMRES method converge without restart at certain iteration steps (on 8.1 of **Section 1.8**),
- (b) some will have no progress with restart at certain iteration steps (on 8.2 and 8.3 of **Section 1.8**), and
- (c) some will have better results with restarting GMRES (**Section 1.9**).

When an initial guess has a slow convergence rate, we will simply choose a new initial guess. An interesting question is how to determine how big “m” should be for our initial approximation. Some of these questions will be answered in another study by the author.

**APPENDIX A:**

$\mathbf{H}_m =$

$$\begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} & \vdots & \vdots & \vdots & \vdots & h_{1,m-2} & h_{1,m-1} & h_{1,m} \\ h_{2,1} & h_{2,2} & h_{2,3} & \vdots & \vdots & \vdots & \vdots & h_{2,m-2} & h_{2,m-1} & h_{2,m} \\ 0 & h_{3,2} & h_{3,3} & \vdots & \vdots & \vdots & \vdots & h_{3,m-2} & h_{3,m-1} & h_{3,m} \\ 0 & 0 & h_{4,3} & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & 0 & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & h_{m-2,m-2} & h_{m-2,m-1} & h_{m-2,m} \\ 0 & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & h_{m-1,m-2} & h_{m-1,m-1} & h_{m-1,m} \\ 0 & 0 & 0 & \vdots & \vdots & \vdots & \vdots & 0 & h_{m,m-1} & h_{m,m} \end{bmatrix}$$

$\mathbf{H}_m^g =$

$$\begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} & \vdots & \vdots & \vdots & \vdots & h_{1,m-2} & h_{1,m-1} & h_{1,m} \\ h_{2,1} & h_{2,2} & h_{2,3} & \vdots & \vdots & \vdots & \vdots & h_{2,m-1} & h_{2,m-1} & h_{2,m} \\ 0 & h_{3,2} & h_{3,3} & \vdots & \vdots & \vdots & \vdots & h_{3,m-2} & h_{3,m-1} & h_{3,m} \\ 0 & 0 & h_{4,3} & \vdots & \vdots & \vdots & \vdots & h_{4,m-2} & h_{4,m-1} & h_{4,m} \\ \vdots & 0 & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & h_{m-1,m-2} & h_{m-1,m-1} & h_{m-1,m} \\ 0 & 0 & 0 & \vdots & \vdots & \vdots & \vdots & 0 & h_{m,m-1} & h_{m,m} \\ 0 & 0 & 0 & \vdots & \vdots & \vdots & \vdots & 0 & 0 & h_{m+1,m} \end{bmatrix}$$

**APPENDIX B:**

$r_m^B \neq 0$  if and only if  $r_m^A \neq 0$

**Proof:** (a) Suppose that  $r_m^B \neq 0$ .

Since  $r_m^B = (r_0 - Az_m^B)$  where  $Az_m^B \in AK_m$ ,

$r_0 \notin AK_m$ .

Since  $r_m^A = (r_0 - Az_m^A)$  where  $Az_m^A \in AK_m$ ,

$r_m^A \neq 0$ . (figure 1)

(b) Conversely, Suppose that  $r_m^A \neq 0$ .

Since  $r_m^A = (r_0 - Az_m^A)$  where  $Az_m^A \in AK_m$ ,

$r_0 \notin AK_m$ .

Since  $r_m^B = (r_0 - Az_m^B)$  where  $Az_m^B \in AK_m$ ,

$r_m^B \neq 0$ . (figure 1)

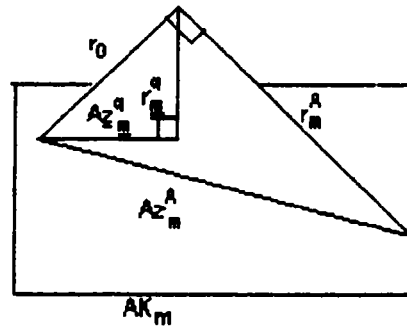


figure 1.

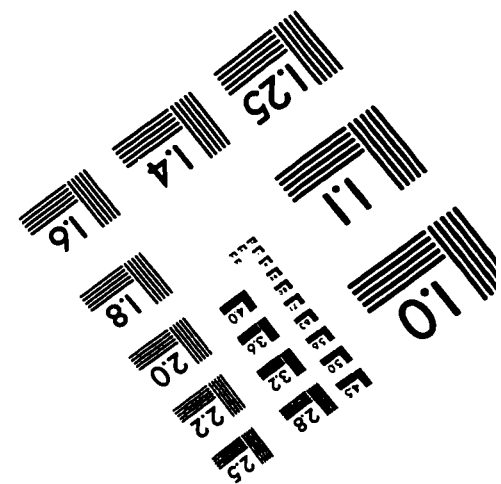
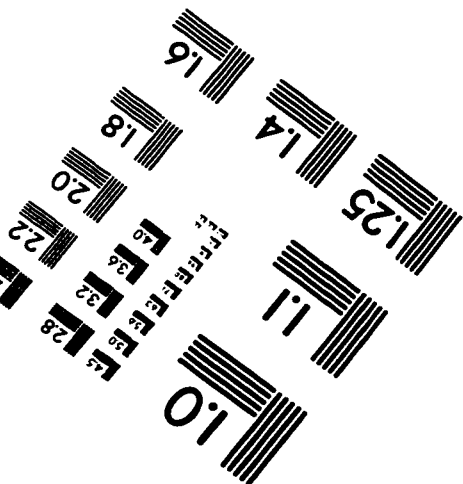
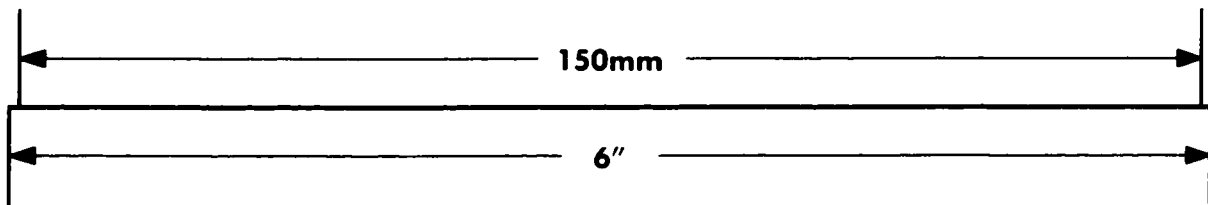
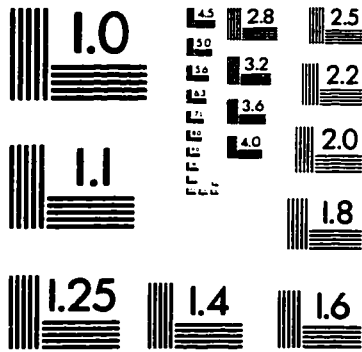
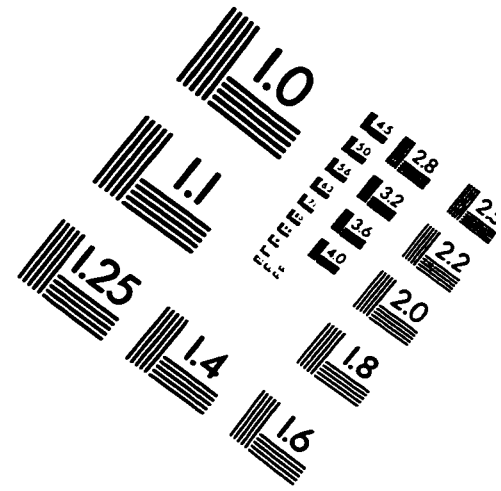
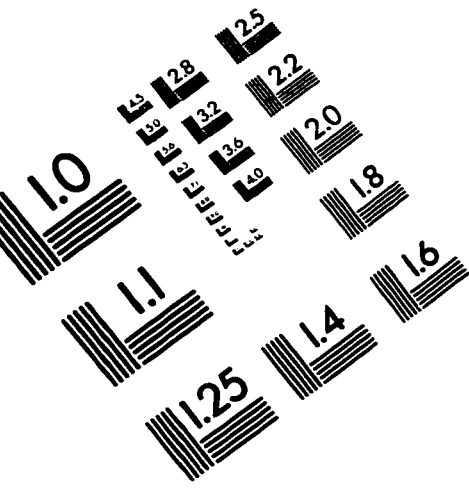


## **Bibliography**

- [1] GMRES : A generalized minimal residual algorithm for solving non-symmetric linear systems. *Youseff Saad and Martin H. Schultz*  
SIAM J. sci. stat. Comput. Vol. 7, No. 3, July 1986.
- [2] A theoretical comparison of the Arnoldi and GMRES algorithms. *Peter N. Brown*  
SIAM J. sci. stat. Comput. Vol. 12, No. 1, pp. 58-78, January 1991.
- [3] Wavelet Transforms and Orthonormal Wavelet Bases. Section 2.3. Orthonormal wavelet bases : The Haar basis as an example. *Ingrid Daubechies*  
Proceedings of symposia in applied Mathematics. Vol. 47. Different perspectives on wavelets. *American Mathematical Society short course.*  
January 11-12, 1993. San Antonio, Texas.
- [4] Krylov subspace Methods for solving unsymmetric linear systems. *Y. Saad*  
Math. Comp., 37 (1981), pp. 105-126.
- [5] Linear Algebra and Its Applications. Addison-Wesley Publishing Company.  
*David C. Lay.* 1994.
- [6] Wavelets and Adapted Waveform Analysis: A tool kit for signal processing and numerical analysis. *Ronald R. Coifman and M. Victor Wickerhauser.*  
Proceedings of Symposia in Applied Mathematics. Vol. 47, 1993.
- [7] Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data. *David L. Donoho.*  
Proceedings of Symposia in Applied Mathematics. Vol. 47, 1993.
- [8] Multiresolution analysis and supercompact multiwavelets.  
*Richard M. Beam and Robert F. Warming.*  
In SIAM annual meeting, Kansas City, Missouri, July 22-26,1996 and  
in NASA TM - 110405, July 1996.
- [9] Surfing the wavelets. *Joshua Altmann.* Internet links. 1997.
- [10] A gentle introduction to wavelets. *E. Rehmi Post.*  
University of Massachusetts, Amherst. Math 442, May 1995.

- [11] **Multidimensional Wavelets and their Applications.** *Eugene Garry Ionel.*  
San Jose State University. 1994.
- [12] **Numerical Linear Algebra and Applications.**  
Brooks/Cole Publishing Company. *Biswa Nath Datta.* 1995.
- [13] **Iterative Methods for Sparse Linear Systems.** PWS Publishing Company.  
*Yousef Saad.* 1996.
- [14] **An Introduction to Wavelets.** *Amara Graps.* *Internet links.* 12 November 1995.
- [15] **Wavelets and Adapted Waveform Analysis.**  
**A Toolkit for Signal Processing and Numerical Analysis.**  
*Ronald R. Coifman and M. Victor Wickerhauser.*  
Proceedings of Symposia in Applied Mathematics, Volume 47, 1993.
- [16] **Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data.** *David L. Donoho.*  
Proceedings of Symposia in Applied Mathematics, Volume 47, 1993.
- [17] **Templates for the Solution of Linear systems: Building Blocks for Iterative Methods.**  
*R. Barrett and M. Berry and T. F. Chan and J. Demmel and J. Donato and J. Dongarra and V. Eijkhout and R. Pozo and C. Romine and H. Van der Vorst.* 1994.

# IMAGE EVALUATION TEST TARGET (QA-3)



**APPLIED IMAGE, Inc**  
 1653 East Main Street  
 Rochester, NY 14609 USA  
 Phone: 716/482-0300  
 Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved