

## Chapter 5

# Krylov Subspace Methods that Are Based on the Minimization of the Residual

**Remark 5.1** *Goal.* The goal of these methods consists in determining

$$\mathbf{x}^{(k)} \in \mathbf{x}^{(0)} + K_k(\mathbf{r}^{(0)}, A)$$

such that the corresponding Euclidean norm of the residual

$$\|\mathbf{r}^{(k)}\|_2 = \|\mathbf{b} - A\mathbf{x}^{(k)}\|_2$$

becomes minimal in the space  $K_k(\mathbf{r}^{(0)}, A)$ . Note, in the Richardson iteration with the special choice (4.4), the norm of the residual will be minimized on the line  $\mathbf{x}^{(k)} + \tau\mathbf{r}^{(k)}$ . However, the minimum on this line is in general not the global minimum in  $\mathbf{x}^{(0)} + K_k(\mathbf{r}^{(0)}, A)$ .  $\square$

### 5.1 General matrices

**Remark 5.2** *Construction of an orthonormal basis of the Krylov subspace.* To perform the minimization of the Euclidean norm of the residual efficiently, an orthonormal basis of  $K_m(\mathbf{q}_1, A)$  is needed. There are several ways to transform an arbitrary basis into an orthonormal one, e.g. the modified Gram<sup>1</sup>–Schmidt<sup>2</sup> method<sup>3</sup> or the Householder<sup>4</sup> algorithm. In the context of Krylov subspace methods, the computation of an orthonormal basis of  $K_m(\mathbf{q}_1, A)$  is called Arnoldi's<sup>5</sup> method.  $\square$

<sup>1</sup>Jorgen Pedersen Gram (1850 – 1916)

<sup>2</sup>Erhard Schmidt (1876 – 1959)

<sup>3</sup>Given a set of orthonormal vectors  $\{\mathbf{u}_1, \dots, \mathbf{u}_{m-1}\}$  and a vector  $\mathbf{v}_m$  that should be orthonormalized with this set. In the original Gram–Schmidt method, one computes all projections of  $\mathbf{v}_m$  with respect to  $\mathbf{u}_i$ ,  $i = 1, \dots, m-1$ , and subtracts these projections. In the modified Gram–Schmidt method, one computes the projection of  $\mathbf{v}_m$  to one of the vectors, say  $\mathbf{u}_1$ , and subtracts this projection. The result  $\mathbf{v}_m^{(1)}$  is orthogonal to  $\mathbf{u}_1$ . Now the projection  $\mathbf{v}_m^{(1)}$  with respect to a second vector, say  $\mathbf{u}_2$  is computed and subtracted to give  $\mathbf{v}_m^{(2)}$ . This vector is orthogonal to  $\mathbf{u}_1$  and  $\mathbf{u}_2$ . Continuing this procedure leads to the orthonormalization of  $\mathbf{v}_m$  with respect to  $\{\mathbf{u}_1, \dots, \mathbf{u}_{m-1}\}$ . In exact arithmetic, both versions are identical. From the numerical point of view, the original Gram–Schmidt method might be unstable whereas the modified Gram–Schmidt method is stable.

<sup>4</sup>Alston Scott Householder (1904 – 1993)

<sup>5</sup>Walter Edwin Arnoldi (1917 – 1995)

**Algorithm 5.3 Arnoldi method, modified Gram–Schmidt method.** Given  $A \in \mathbb{R}^{n \times n}$  and  $\mathbf{q}_1 \in \mathbb{R}^n$  with  $\|\mathbf{q}_1\|_2 = 1$ .

```

1. for  $j = 1 : m$ 
2.    $\mathbf{w}_j = A\mathbf{q}_j$ 
3.   for  $i = 1 : j$ 
4.      $h_{ij} = (\mathbf{w}_j, \mathbf{q}_i)$ 
5.      $\mathbf{w}_j = \mathbf{w}_j - h_{ij}\mathbf{q}_i$            % subtract projection
6.   endfor
7.    $h_{j+1,j} = \|\mathbf{w}_j\|_2$ 
8.   if  $h_{j+1,j} == 0$ 
9.     stop
10.  endif
11.   $\mathbf{q}_{j+1} = \mathbf{w}_j/h_{j+1,j}$ 
12. endfor

```

□

**Lemma 5.4 Computation of an orthonormal basis by Arnoldi’s method.** If  $\dim K_m(\mathbf{q}_1, A) = m$ , then Arnoldi’s method computes an orthonormal basis  $\{\mathbf{q}_1, \dots, \mathbf{q}_m\}$  of  $K_m(\mathbf{q}_1, A)$ .

**Proof:** The vectors  $\mathbf{q}_1, \dots, \mathbf{q}_m$  are orthonormal by construction: orthogonal by line 3 – 6 and normalized by line 7 and 11. One has to show that they belong all to  $K_m(\mathbf{q}_1, A)$ . This statement will follow from the fact that each vector  $\mathbf{q}_j$  is of the form  $p_{j-1}(A)\mathbf{q}_1$ , where  $p_{j-1}$  is a polynomial of degree  $j - 1$ . The proof is done by induction. For  $j = 1$ , one has  $\mathbf{q}_1 = p_0(A)\mathbf{q}_1$  such that  $p_0(t) = 1$ . Assume, the statement is true for  $j = k$ . One has by using first line 11, then lines 3 – 6, and finally the assumption of the induction

$$h_{k+1,k}\mathbf{q}_{k+1} = \mathbf{w}_k = A\mathbf{q}_k - \sum_{i=1}^k h_{ik}\mathbf{q}_i = Ap_{k-1}(A)\mathbf{q}_1 - \sum_{i=1}^k h_{ik}p_{i-1}(A)\mathbf{q}_1 = p_k(A)\mathbf{q}_1.$$

■

**Remark 5.5 Factorization of the system matrix.** Denote

$$Q_m = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m) \in \mathbb{R}^{n \times m},$$

$$H_m = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1,m-1} & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2,m-1} & h_{2m} \\ 0 & h_{32} & \cdots & h_{3,m-1} & h_{3m} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{m,m-1} & h_{mm} \\ 0 & 0 & \cdots & 0 & h_{m+1,m} \end{pmatrix} \in \mathbb{R}^{(m+1) \times m}.$$

A matrix of this form, i.e.,  $h_{ij} = 0$  for  $i > j + 1$ , is called (upper) Hessenberg<sup>6</sup> matrix. It follows readily from Arnoldi’s method that

$$AQ_m = Q_{m+1}H_m. \tag{5.1}$$

□

**Remark 5.6 Initial vector in Krylov subspace methods.** In the Krylov subspace methods,  $\mathbf{r}^{(0)}/\|\mathbf{r}^{(0)}\|_2$  plays the role of  $\mathbf{q}_1$  in Arnoldi’s method. □

---

<sup>6</sup>Karl Hessenberg (1904 – 1959)

**Remark 5.7** *Principle approach for minimizing the residual.* The goal of the methods presented in this section is to minimize the Euclidean norm of the residual. One has

$$\begin{aligned}\|\mathbf{r}^{(k)}\|_2 &= \|\mathbf{b} - A\mathbf{x}^{(k)}\|_2 = \|\mathbf{r}^{(0)} + A\mathbf{x}^{(0)} - A\mathbf{x}^{(k)}\|_2 \\ &= \|\mathbf{r}^{(0)} - A(\mathbf{x}^{(k)} - \mathbf{x}^{(0)})\|_2\end{aligned}$$

with  $\mathbf{x}^{(k)} - \mathbf{x}^{(0)} \in K_k(\mathbf{r}^{(0)}, A)$ , see Remark 4.8. Since the vectors  $\{\mathbf{q}_1, \dots, \mathbf{q}_k\}$  computed with Arnoldi's method form a generating system of  $K_k(\mathbf{r}^{(0)}, A)$ , it is

$$\mathbf{x}^{(k)} - \mathbf{x}^{(0)} = \sum_{i=1}^k z_i \mathbf{q}_i = Q_k \mathbf{z}$$

with  $\mathbf{z} = (z_1, \dots, z_k)^T$ ,  $Q_k = (\mathbf{q}_1, \dots, \mathbf{q}_k)$ . Using (5.1),  $Q_{k+1} \mathbf{e}_1 = \mathbf{q}_1 = \mathbf{r}^{(0)} / \|\mathbf{r}^{(0)}\|_2$  and the fact that the Euclidean norm is invariant under orthonormal transformations, one obtains

$$\begin{aligned}\|\mathbf{r}^{(k)}\|_2^2 &= \|\mathbf{r}^{(0)} - AQ_k \mathbf{z}\|_2^2 = \|\mathbf{r}^{(0)} - Q_{k+1} H_k \mathbf{z}\|_2^2 \\ &= \left\| \|\mathbf{r}^{(0)}\|_2 Q_{k+1} \mathbf{e}_1 - Q_{k+1} H_k \mathbf{z} \right\|_2^2 = \left\| Q_{k+1} \left( \|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1 - H_k \mathbf{z} \right) \right\|_2^2 \\ &= \left\| \|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1 - H_k \mathbf{z} \right\|_2^2.\end{aligned}$$

The minimizer of the residual is obtained by solving the least squares problem

$$\min_{\mathbf{z} \in \mathbb{R}^k} \left\| \|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1 - H_k \mathbf{z} \right\|_2^2. \quad (5.2)$$

This problem possesses  $k$  unknowns and the vector that has to be minimized has  $k+1$  components. It can be solved, e.g., with the QR algorithm, see lecture notes Numerical Mathematics I. Let  $\mathbf{z}^{(k)}$  be a solution of this problem, then the next iterate of the Krylov subspace method is

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k \mathbf{z}^{(k)}. \quad (5.3)$$

This algorithm is called GMRES (generalized minimal residual). It has been proposed the first time in Saad and Schultz (1986).  $\square$

### Theorem 5.8 Properties of GMRES.

- i) In the case that Arnoldi's method has an early breakdown, i.e.,  $h_{l+1,l} = 0$ , then  $\dim K_k(\mathbf{r}^{(0)}, A) = l < k$  and  $\mathbf{r}^{(l)} = \mathbf{0}$ . Hence  $A\mathbf{x}^{(l)} = \mathbf{b}$ .*
- ii) The iterate  $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k \mathbf{z}^{(k)}$  is uniquely determined.*
- iii) It holds*

$$\|\mathbf{r}^{(k)}\|_2 \leq \|\mathbf{r}^{(k-1)}\|_2, \quad k = 1, 2, 3, \dots$$

**Proof:** *i)* The breakdown of Arnoldi's method after  $l$  steps,  $h_{l+1,l} = 0$ , is equivalent to  $\mathbf{w}_l = \mathbf{0}$ . It follows from Arnoldi's algorithm lines 3 – 6 that

$$A\mathbf{q}_l = \sum_{i=1}^l h_{il} \mathbf{q}_i,$$

where  $\mathbf{q}_1 = \mathbf{r}^{(0)} / \|\mathbf{r}^{(0)}\|_2$  in the case of GMRES. Hence, one has  $\dim K_{l+1}(\mathbf{r}^{(0)}, A) = \dim K_l(\mathbf{r}^{(0)}, A)$ . One obtains by induction

$$\dim K_k(\mathbf{r}^{(0)}, A) = \dim K_l(\mathbf{r}^{(0)}, A) \quad \text{for } k \geq l.$$

Using matrix notations, Arnoldi's method gives in this case

$$AQ_l = Q_l \tilde{H}_l \quad \text{with} \quad \tilde{H}_l = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1l} \\ h_{21} & h_{22} & \cdots & h_{2l} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & h_{l,l-1} & h_{ll} \end{pmatrix} \in \mathbb{R}^{l \times l}, \quad Q_l \in \mathbb{R}^{n \times l}.$$

Since  $A$  is non-singular and  $\text{rank}(Q_l) = l$ , one has  $\text{rank}(AQ_l) = l$ . Consequently, it is  $\text{rank}(Q_l \tilde{H}_l) = l$  and  $\text{rank}(\tilde{H}_l) = l$  and  $\tilde{H}_l$  is invertible. In the same way as above, one obtains

$$\min \|\mathbf{r}^{(l)}\|_2^2 = \min_{\mathbf{z} \in \mathbb{R}^l} \left\| \|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1 - \tilde{H}_l \mathbf{z} \right\|_2^2.$$

The minimizer is given by  $\mathbf{z}^{(l)} = \tilde{H}_l^{-1} \left( \|\mathbf{r}^{(0)}\|_2 \mathbf{e}_1 \right)$  which gives  $\|\mathbf{r}^{(l)}\|_2 = 0$ .

*ii)* If  $\text{rank}(H_k) = k$ , the minimizer of (5.2) is unique (theory of least squares problems, see Numerical Mathematics I). If  $\text{rank}(H_k) < k$ , then  $\mathbf{x}^{(k)} = \mathbf{x}$ , see i).

*iii)* The set in which the minimizer is computed becomes larger since the inclusion  $K_k(\mathbf{r}^{(0)}, A) \supseteq K_{k-1}(\mathbf{r}^{(0)}, A)$  holds.  $\blacksquare$

**Remark 5.9** *Implementational issues.*

- The GMRES process consists in principle of two steps:
  1. computing the orthonormal basis of  $K_k(\mathbf{r}^{(0)}, A)$ ,
  2. solving the least squares problem (5.2) to find the minimizer of the residual with a standard method.

In the practical use of GMRES, Step 2 is performed only at the end of the iteration. Thus, the iterate  $\mathbf{x}^{(k)}$  is not directly available. It is computed in a post-processing step. However, there is an elegant and inexpensive way to compute  $\|\mathbf{r}^{(k)}\|_2$  without having access to  $\mathbf{x}^{(k)}$ , see Saad (2003). With  $\|\mathbf{r}^{(k)}\|_2$  one can control the iterative process.

For concrete ways to implement GMRES, it is referred to Saad and Schultz (1986); Saad (2003).

- Each step of GMRES requires one matrix-vector multiplication, line 2 of Arnoldi's method.
- In exact arithmetic, GMRES terminates with the solution in at most  $n$  steps. This property is, however, of no practical use for large  $n$ .
- From the practical point of view, the greatest problem of GMRES is that one has to store the basis  $\{\mathbf{q}_1, \dots, \mathbf{q}_k\}$  of  $K_k(\mathbf{r}^{(0)}, A)$ . see lines 3 – 6 of Arnoldi's method. Thus, with every new iteration, one has to store an additional vector. This situation is called long recurrence. In practice, one prescribes a maximal order  $m$  of the Krylov subspace. After  $m$  iterations, GMRES is stopped with the iterate  $\mathbf{x}^{(m)}$ . If  $\mathbf{x}^{(m)}$  is not yet sufficiently close to the solution, GMRES is started from the beginning with  $\mathbf{x}^{(0)} = \mathbf{x}^{(m)}$ . This approach is called GMRES(m) (with restart). An optimal choice of  $m$  is in general an unresolved problem. Often  $m \in [5, 20]$  is used. GMRES(m) might also fail to converge, see Saad and Schultz (1986) for the simple example

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

GMRES converges in two steps whereas GMRES(1) computes the stationary sequence  $\mathbf{x}^{(1)} = \mathbf{x}^{(0)}$ ,  $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} = \mathbf{x}^{(0)}$  and so on. Despite of the possibility to fail, GMRES(m) is one of the most popular and best performing iterative methods for solving linear systems of equations with non-symmetric matrix.  $\square$

## 5.2 Symmetric Matrices

**Remark 5.10** *Goal.* Arnoldi's method and the minimization of the residual in  $K_k(\mathbf{r}^{(0)}, A)$  will be studied in the special case that  $A$  is symmetric. The most important result will be that in this case, it is not necessary to store the basis of  $K_k(\mathbf{r}^{(0)}, A)$ . It suffices to store a fixed number of only few basis vectors. Thus, the memory requirements do not increase in the course of the iteration and the most important problem of using GMRES vanishes.  $\square$

**Remark 5.11** *Arnoldi's method revisited.* First, Arnoldi's method is revisited. From the general relation (5.1) it follows by the orthonormality of the columns of  $Q_k$  and  $Q_{k+1}$  that

$$Q_k^T A Q_k = Q_k^T Q_{k+1} H_k = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} H_k =: \tilde{H}_k \in \mathbb{R}^{k \times k}. \quad (5.4)$$

Thus,  $\tilde{H}_k$  contains just the first  $k$  rows of  $H_k$ . Since

$$(Q_k^T A Q_k)^T = Q_k^T A^T Q_k = Q_k^T A Q_k,$$

is a symmetric matrix,  $\tilde{H}_k$  is symmetric, too. As in the case of a general matrix,  $\tilde{H}_k$  is an upper Hessenberg matrix. From its symmetry, it follows that  $\tilde{H}_k$  is even a tridiagonal matrix. Hence,  $H_k$  is a tridiagonal matrix, too

$$H_k = \begin{pmatrix} \alpha_1 & \beta_2 & 0 & \cdots & 0 & 0 \\ \beta_2 & \alpha_2 & \beta_3 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_{k-1} & \beta_k \\ 0 & 0 & 0 & \cdots & \beta_k & \alpha_k \\ 0 & 0 & 0 & \cdots & 0 & \beta_{k+1} \end{pmatrix} \in \mathbb{R}^{(k+1) \times k}.$$

Arnoldi's method simplifies. Using (5.1) and the special form of  $H_k$ , one obtains the relation

$$A \mathbf{q}_k = \beta_k \mathbf{q}_{k-1} + \alpha_k \mathbf{q}_k + \beta_{k+1} \mathbf{q}_{k+1}.$$

From this relation,  $\mathbf{q}_{k+1}$  can be computed. The corresponding algorithm is called Lanczos<sup>7</sup> algorithm.  $\square$

**Algorithm 5.12 Lanczos algorithm – modified Gram–Schmidt variant.**

Given a symmetric matrix  $A \in \mathbb{R}^{n \times n}$  and  $\mathbf{q}_1 \in \mathbb{R}^n$  with  $\|\mathbf{q}_1\|_2 = 1$ .

1.  $\beta_1 = 0$
2.  $\mathbf{q}_0 = 0$
3. **for**  $j = 1 : m$
4.      $\mathbf{s} = A \mathbf{q}_j - \beta_j \mathbf{q}_{j-1}$
5.      $\alpha_j = (\mathbf{s}, \mathbf{q}_j)$
6.      $\mathbf{s} = \mathbf{s} - \alpha_j \mathbf{q}_j$
7.      $\beta_{j+1} = \|\mathbf{s}\|_2$
8.     **if**  $\beta_{j+1} == 0$
9.         **stop**

---

<sup>7</sup>Cornelius Lanczos (1893 – 1974)



Since  $H_k$  is tridiagonal, one obtains

$$\bar{R}_k = \begin{pmatrix} r_{11} & r_{12} & r_{13} & 0 & \cdots & 0 \\ 0 & r_{22} & r_{23} & r_{24} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & & & & & r_{k-2,k} \\ 0 & & & & & r_{k-1,k} \\ 0 & & & & & r_{k,k} \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

i.e.  $r_{ij} = 0$  if  $j > i + 2$ . A Givens rotation changes only the two columns that are involved, i.e. here the two neighbouring columns  $j$  and  $j + 1$ . Since the non-zero entries at column  $j$  of  $H_k$  are at rows  $(j - 1), j$ , and  $j + 1$ , where the last will be transformed to become zero, a new non-zero entry at column  $(j + 1)$  can occur only at row  $(j - 1)$ .

Consider the only interesting case  $\mathbf{r}^{(k)} \neq \mathbf{0}$ , in which the matrix  $H_k$  has full rank. Let  $R_k$  be the matrix which consists of the first  $k$  rows of  $\bar{R}_k$ . The matrix  $R_k$  is non-singular since  $H_k$  has full rank. Setting

$$P_k = (\mathbf{p}_1 \mathbf{p}_2 \dots \mathbf{p}_k) := Q_k R_k^{-1} \in \mathbb{R}^{n \times k},$$

one has from  $P_k R_k = Q_k$  and due to the special form of  $R_k$  the recursion

$$\begin{aligned} \mathbf{p}_1 &= \frac{\mathbf{q}_1}{r_{11}}, \\ \mathbf{p}_2 &= \frac{1}{r_{22}} (\mathbf{q}_2 - r_{12} \mathbf{p}_1) \quad (\Leftarrow r_{22} \mathbf{p}_2 + r_{12} \mathbf{p}_1 = \mathbf{q}_2), \\ &\vdots \\ \mathbf{p}_j &= \frac{1}{r_{jj}} (\mathbf{q}_j - \mathbf{p}_{j-1} r_{j-1,j} - \mathbf{p}_{j-2} r_{j-2,j}), \quad j = 3, \dots, k. \end{aligned} \quad (5.5)$$

The least squares problem (5.2) can now rewritten in the form

$$\min_{\mathbf{z} \in \mathbb{R}^k} \left\| \left\| \mathbf{r}^{(0)} \right\|_2 \mathbf{e}_1 - G_1 G_2 \cdots G_k \bar{R}_k \mathbf{z} \right\|_2^2 = \min_{\mathbf{z} \in \mathbb{R}^k} \left\| \left\| \mathbf{r}^{(0)} \right\|_2 G_k^T \cdots G_2^T G_1^T \mathbf{e}_1 - \bar{R}_k \mathbf{z} \right\|_2^2,$$

because the Euclidean norm is invariant under a multiplication with a unitary matrix. Since the last row of  $\bar{R}_k$  vanishes, its Moore<sup>9</sup>–Penrose<sup>10</sup> inverse (pseudo-invers), see Numerical Mathematics I, is given by

$$\bar{R}_k^+ = (R_k^{-1} \mathbf{0}) \in \mathbb{R}^{k \times (k+1)}$$

and the solution of the least squares problem is given by

$$\mathbf{z}^{(k)} = \left\| \left\| \mathbf{r}^{(0)} \right\|_2 \bar{R}_k^+ G_k^T \cdots G_2^T G_1^T \mathbf{e}_1 \right\|_2 = \left\| \left\| \mathbf{r}^{(0)} \right\|_2 R_k^{-1} (G_k^T \cdots G_1^T \mathbf{e}_1) \right\|_{1 \leq i \leq k},$$

where the last index symbolizes that only the first  $k$  components of the vectors are taken. Consequently, the iterate with the minimal residual has the form, see (5.3),

$$\begin{aligned} \mathbf{x}^{(k)} &= \mathbf{x}^{(0)} + Q_k \mathbf{z}^{(k)} = \mathbf{x}^{(0)} + \left\| \left\| \mathbf{r}^{(0)} \right\|_2 Q_k R_k^{-1} \underbrace{(G_k^T \cdots G_1^T \mathbf{e}_1)}_{\in \mathbb{R}^{k+1}} \right\|_{1 \leq i \leq k}, \\ &= \mathbf{x}^{(0)} + \left\| \left\| \mathbf{r}^{(0)} \right\|_2 P_k (G_k^T \cdots G_1^T \mathbf{e}_1) \right\|_{1 \leq i \leq k}. \end{aligned}$$

<sup>9</sup>Eliakim Hastings Moore (1862 – 1932)

<sup>10</sup>Roger Penrose, born 1931

Since the Givens rotation or reflection  $G_j^T$  influences only the rows  $j$  and  $j+1$  of the vector to which it is applied, the first  $(j-1)$  of its components stay unchanged:

$$(G_j^T \cdots G_1^T \mathbf{e}_1)_{1 \leq i \leq j-1} = (G_{j-1}^T \cdots G_1^T \mathbf{e}_1)_{1 \leq i \leq j-1}.$$

It follows that

$$\begin{aligned} \mathbf{x}^{(k)} &= \mathbf{x}^{(0)} + \left\| \mathbf{r}^{(0)} \right\|_2 P_{k-1} (G_k^T \cdots G_1^T \mathbf{e}_1)_{1 \leq i \leq k-1} + \left\| \mathbf{r}^{(0)} \right\|_2 \mathbf{p}_k (G_k^T \cdots G_1^T \mathbf{e}_1)_{i=k} \\ &= \mathbf{x}^{(0)} + \underbrace{\left\| \mathbf{r}^{(0)} \right\|_2 P_{k-1} (G_{k-1}^T \cdots G_1^T \mathbf{e}_1)_{1 \leq i \leq k-1}}_{=\mathbf{x}^{(k-1)}} + \left\| \mathbf{r}^{(0)} \right\|_2 \mathbf{p}_k (G_k^T \cdots G_1^T \mathbf{e}_1)_{i=k} \\ &= \mathbf{x}^{(k-1)} + \left\| \mathbf{r}^{(0)} \right\|_2 (G_k^T \cdots G_1^T \mathbf{e}_1)_{i=k} \mathbf{p}_k. \end{aligned}$$

For computing  $\mathbf{p}_k$ , one needs, see (5.5),  $\mathbf{q}_k$ ,  $\mathbf{p}_{k-1}$ , and  $\mathbf{p}_{k-2}$ . The result  $\mathbf{p}_k$  can be stored in place of  $\mathbf{p}_{k-2}$  since  $\mathbf{p}_{k-2}$  is not needed any longer. Together with the short recurrence of the Lanczos algorithm, it is shown that the storage of the basis of  $K_k(\mathbf{r}^{(0)}, A)$  is not necessary.

The resulting method which computes iterates with minimal residual for symmetric matrices  $A$  is called MINRES. MINRES requires to store six arrays:  $\mathbf{q}_k$ ,  $\mathbf{q}_{k+1}$ ,  $\mathbf{s}$ ,  $\mathbf{p}_k$ ,  $\mathbf{p}_{k-1}$ , and  $\mathbf{x}^{(k)}$ . In comparison with GMRES, the current iterate  $\mathbf{x}^{(k)}$  is known and not only the residual of the current iterate.  $\square$

**Remark 5.16** *S.p.d. matrices, conjugate residual method.* In practice,  $A$  is often not only symmetric but also positive definite. In this case, MINRES is seldom used, since for s.p.d. matrices there is a more efficient method called conjugate residual method.  $\square$

**Definition 5.17 Conjugate vectors.** Let  $A \in \mathbb{R}^{n \times n}$  be symmetric and positive definite. The vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  are called ( $A$ )-orthogonal or ( $A$ )-conjugate if

$$\mathbf{x}^T A \mathbf{y} = (A\mathbf{x}, \mathbf{y}) = 0.$$

If there is no ambiguity, the vectors are called just conjugate.  $\square$

**Remark 5.18** *Comparison of MINRES and conjugate residuals, conjugate gradient method.* The conjugate residual method needs to store only five arrays. It requires in each iteration one matrix-vector product. The memory requirements are one array more than the conjugate gradient method, see Section 6.2. In addition, one has to compute one vector update ( $2n$  flops) per iteration more with the conjugate residual method in comparison with the conjugate gradient method. Since both methods need in general a similar number of iterations, the conjugate gradient method is preferred in practice. For this reason, it will be referred to the literature for more details concerning the conjugate residual method.  $\square$

**Remark 5.19** *S.p.d. matrices vs. other matrices.* As it can be already seen in the case of Krylov subspace methods that minimize the residual, one has to distinguish the cases that  $A$  is s.p.d. or  $A$  is another matrix. These two cases represent two worlds in the context of iterative methods for solving linear systems of equations. Methods that can be used for general matrices and that are considered to work usually well in this case, are generally not the best methods for s.p.d. matrices. The solution of systems with s.p.d. matrices is much simpler. In engineering practice, it is a common approach to try to reduce the solution of a complicated problem to the successive solution of linear systems with s.p.d. matrices.  $\square$