Chapter 9
Other Krylov Subspace Methods for Non-Symmetric Systems

Remark 9.1. Motivation. The Krylov subspace methods GMRES and FOM for solving general linear systems of equations have the disadvantage that their costs (in memory and flops) increases with the number of iterations, since there is no short recurrence. A remedy is to use restarted versions, compare Remark 5.9. However, the restart might lead to a considerably slower rate of convergence. This section presents alternative approaches that are based on short recurrences but which fulfill the properties of minimizing the residual (like GMRES) or of the residual being orthogonal to the Krylov subspace (like FOM), respectively, not longer. □

Remark 9.2. Biorthogonal bases. The starting point of the alternative algorithms is the construction of a pair of biorthogonal bases

$$\left\{\underline{v}^{(1)},\ldots,\underline{v}^{(k)}\right\} \text{ of } K_k\left(\underline{r}^{(0)},A\right) = \operatorname{span}\left\{\underline{r}^{(0)},A\underline{r}^{(0)},\ldots,A^{k-1}\underline{r}^{(0)}\right\},$$
$$\left\{\underline{w}^{(1)},\ldots,\underline{w}^{(k)}\right\} \text{ of } K_k\left(\underline{r}^{(0)},A^T\right) = \operatorname{span}\left\{\underline{r}^{(0)},A^T\underline{r}^{(0)},\ldots,\left(A^T\right)^{k-1}\underline{r}^{(0)}\right\}$$

such that

$$\left(\underline{w}^{(j)},\underline{v}^{(i)}\right) = \delta_{ij}.$$

These bases can be constructed with the Lanczos biorthogonalization procedure. □

Algorithm 9.3. Lanczos biorthogonalization procedure. Given a matrix $A \in \mathbb{R}^{n\times n}$ and $\underline{r}^{(0)} \in \mathbb{R}^n$.

1. $\underline{v}^{(1)} = \underline{r}^{(0)}/\left\|\underline{r}^{(0)}\right\|_2$
2. $\underline{w}^{(1)} = \underline{v}^{(1)}$
3. $\beta_0 = 0$, $\gamma_0 = 0$
4. $\underline{v}^{(0)} := \underline{0}$, $\underline{w}^{(0)} = \underline{0}$
5. `for` $j = 1 : k$

6. $\underline{s} = A\underline{v}^{(j)}$
7. $\underline{z} = A^T\underline{w}^{(j)}$
8. $\alpha_j = (\underline{w}^{(j)}, \underline{s})$
9. $\underline{\tilde{v}}^{(j+1)} = \underline{s} - \alpha_j\underline{v}^{(j)} - \beta_{j-1}\underline{v}^{(j-1)}$
10. $\underline{\tilde{w}}^{(j+1)} = \underline{z} - \alpha_j\underline{w}^{(j)} - \gamma_{j-1}\underline{w}^{(j-1)}$
11. $\gamma_j = \left\|\underline{\tilde{v}}^{(j+1)}\right\|_2$
12. $\underline{v}^{(j+1)} = \underline{\tilde{v}}^{(j+1)}/\gamma_j$
13. $\beta_j = \left(\underline{\tilde{w}}^{(j+1)}, \underline{v}^{(j+1)}\right)$
14. $\underline{w}^{(j+1)} = \underline{\tilde{w}}^{(j+1)}/\beta_j$
15. endfor

□

Remark 9.4. On the Lanczos biorthogonalization procedure.

- There is a short recurrence in Algorithm 9.3.
- Note that the basis of $K_k\left(\underline{r}^{(0)}, A\right)$ will be in general not orthogonal as well as the basis of $K_k\left(\underline{r}^{(0)}, A^T\right)$. For non-symmetric matrices, the computation of an orthogonal basis is not possible with a short recurrence.
- In the case $A = A^T$, Algorithm 9.3 is exactly the Lanczos Algorithm 5.12.
- Algorithm 9.3 requires two matrix-vector products, lines 6 and 7.
- A critical point of Algorithm 9.3 is the product of the transposed of A with a vector, line 7. In some applications, A is not given explicitly. Then, A^T is usually not available. But much more important, the application of a number of preconditioners, see Chapter 8, becomes complicated if A^T appears in the algorithm.

□

Theorem 9.5. Computation of a pair of biorthogonal bases. *Assume that $\left(\underline{\tilde{w}}^{(j)}, \underline{v}^{(j)}\right) \neq 0$ for all $j = 1, \cdots, k$. Then, the Lanczos biorthogonalization procedure computes a pair of biorthogonal bases.*

Proof. The theorem is proved by induction. The statement is true if $k = 1$ since $\underline{w}^{(1)} = \underline{v}^{(1)}$, line 2 and $\left\|\underline{v}^{(1)}\right\|_2 = 1$, line 1. For $k = 2$, it can be proved directly from the algorithm, in a similar way as for the general case.

Assume, the statement is proved for $i = 1, \ldots, k-1$ with $k - 1 \geq 2$, and suppose $\left(\underline{\tilde{w}}^{(j)}, \underline{v}^{(i)}\right) = \left(\underline{w}^{(j)}, \underline{v}^{(i)}\right) = 0$ for $i \neq j$, $1 \leq i, j \leq k-1$ and $\left(\underline{w}^{(i)}, \underline{v}^{(i)}\right) = 1$, $1 \leq i \leq k-1$. The choice of γ_{k-1} implies $\left\|\underline{v}^{(k)}\right\|_2 = 1$ and line 14 and the choice of β_{k-1} give

$$\left(\underline{w}^{(k)}, \underline{v}^{(k)}\right) = \left(\frac{\underline{\tilde{w}}^{(k)}}{\left(\underline{\tilde{w}}^{(k)}, \underline{v}^{(k)}\right)}, \underline{v}^{(k)}\right) = \frac{\left(\underline{\tilde{w}}^{(k)}, \underline{v}^{(k)}\right)}{\left(\underline{\tilde{w}}^{(k)}, \underline{v}^{(k)}\right)} = 1.$$

Using lines 12 and 9, the assumption of the induction, and the definition of α_{k-1} leads to

$$\left(\underline{w}^{(k-1)}, \underline{v}^{(k)}\right) = \left(\underline{w}^{(k-1)}, \frac{\underline{\tilde{v}}^{(k)}}{\gamma_{k-1}}\right)$$
$$= \frac{1}{\gamma_{k-1}}\left[\left(\underline{w}^{(k-1)}, A\underline{v}^{(k-1)}\right) - \alpha_{k-1}\underbrace{\left(\underline{w}^{(k-1)}, \underline{v}^{(k-1)}\right)}_{=1} - \beta_{k-2}\underbrace{\left(\underline{w}^{(k-1)}, \underline{v}^{(k-2)}\right)}_{=0}\right] = 0.$$

One obtains analogously $\left(\underline{w}^{(k)}, \underline{v}^{(k-1)}\right) = 0$. Moreover, using the lines 12, 9, the assumption of the induction, line 10 (with $\underline{z} = A^T\underline{w}^{(k-2)}$), and the definition of β_{k-2} gives

$$\left(\underline{w}^{(k-2)}, \underline{v}^{(k)}\right)$$
$$= \frac{1}{\gamma_{k-1}}\left[\left(\underline{w}^{(k-2)}, A\underline{v}^{(k-1)}\right) - \alpha_{k-1}\underbrace{\left(\underline{w}^{(k-2)}, \underline{v}^{(k-1)}\right)}_{=0} - \beta_{k-2}\underbrace{\left(\underline{w}^{(k-2)}, \underline{v}^{(k-2)}\right)}_{=1}\right]$$
$$= \frac{1}{\gamma_{k-1}}\left[\left(\underline{w}^{(k-2)}, A\underline{v}^{(k-1)}\right) - \beta_{k-2}\right]$$
$$= \frac{1}{\gamma_{k-1}}\left[\left(A^T\underline{w}^{(k-2)}, \underline{v}^{(k-1)}\right) - \beta_{k-2}\right]$$
$$= \frac{1}{\gamma_{k-1}}\left[\left(\underline{\tilde{w}}^{(k-1)}, \underline{v}^{(k-1)}\right) + \alpha_{k-2}\underbrace{\left(\underline{\tilde{w}}^{(k-2)}, \underline{v}^{(k-1)}\right)}_{=0} + \gamma_{k-3}\underbrace{\left(\underline{\tilde{w}}^{(k-3)}, \underline{v}^{(k-1)}\right)}_{=0} - \beta_{k-2}\right]$$
$$= 0.$$

Analogously, one checks that $\left(\underline{w}^{(k)}, \underline{v}^{(k-2)}\right) = 0$ and in a similar way, one obtains $\left(\underline{w}^{(k)}, \underline{v}^{(j)}\right) = 0$ and $\left(\underline{w}^{(j)}, \underline{v}^{(k)}\right) = 0$ for $j < k-2$. ∎

Remark 9.6. Matrix representation of the Lanczos biorthogonalization procedure. For the matrix representation of the Lanczos biorthogonalization procedure, the matrices

$$V_k = \left(\underline{v}^{(1)} \ldots \underline{v}^{(k)}\right), \; W_k = \left(\underline{w}^{(1)} \ldots \underline{w}^{(k)}\right) \in \mathbb{R}^{n\times k}$$

are introduced. From Algorithm 9.3, it follows that

$$AV_k = V_{k+1}T_{k+1,k}, \quad A^T W_k = W_{k+1}\hat{T}_{k+1,k} \tag{9.1}$$

with

$$T_{k+1,k} = \begin{pmatrix} \alpha_1 & \beta_1 & & & \\ \gamma_1 & \alpha_2 & & & \\ & \ddots & \ddots & \ddots & \\ & & & & \beta_{k-1} \\ & & & & \alpha_k \\ & & & & \gamma_k \end{pmatrix}, \; \hat{T}_{k+1,k} = \begin{pmatrix} \alpha_1 & \gamma_1 & & & \\ \beta_1 & \alpha_2 & & & \\ & \ddots & \ddots & \ddots & \\ & & & & \gamma_{k-1} \\ & & & & \alpha_k \\ & & & & \beta_k \end{pmatrix},$$