Chapter 7
Convergence of Krylov Subspace Methods

Remark 7.1. Motivation. The Krylov subspace methods compute the solution of (1.1) in at most n iterations (in exact arithmetic) by construction. However, this property is useless if n is large. The question arises whether one can get information about the iterate $\underline{x}^{(k)}$ for $k < n$. $\Box$

Remark 7.2. Starting point of the convergence analysis. The basis of the convergence analysis for Krylov subspace methods is the following observation: $\underline{z} \in K_k\left(\underline{r}^{(0)}, A\right)$ is equivalent to $\underline{z} = q_{k-1}(A)\,\underline{r}^{(0)}$, where $q_{k-1} \in P_{k-1}$ is a polynomial of degree $k-1$. It follows for the residual of the k-th iterate that

$$\begin{aligned} \underline{r}^{(k)} &= \underline{b} - A\underline{x}^{(k)} = \underline{b} - A\left(\underline{x}^{(0)} + \underline{z}\right) = \underline{r}^{(0)} - A\underline{z} = \underline{r}^{(0)} - Aq_{k-1}(A)\,\underline{r}^{(0)} \\ &= p_k(A)\underline{r}^{(0)}, \end{aligned} \tag{7.1}$$

where $p_k(x) = 1 - xq_{k-1}(x) \in P_k$, so that $p_k(0) = 1$.

Considering the methods that are based on the minimization of the residual, see Chapter 5, one has now

$$\left\|\underline{r}^{(k)}\right\|_2 = \min_{p_k \in P_k, p_k(0)=1} \left\|p_k(A)\underline{r}^{(0)}\right\|_2,$$

such that with $\left\|p_k(A)\underline{r}^{(0)}\right\|_2 \le \|p_k(A)\|_2 \left\|\underline{r}^{(0)}\right\|_2$, it follows that

$$\frac{\left\|\underline{r}^{(k)}\right\|_2}{\left\|\underline{r}^{(0)}\right\|_2} \le \min_{p_k \in P_k, p_k(0)=1} \|p_k(A)\|_2\,. \tag{7.2}$$

For all Krylov subspace methods, in particular for those methods which are based on projecting the residual, see Chapter 6, it holds with (7.1) that

$$\underline{x} - \underline{x}^{(k)} = A^{-1}\underline{b} - A^{-1}\left(\underline{b} - \underline{r}^{(k)}\right) = A^{-1}\underline{r}^{(k)} = A^{-1}p_k(A)\,\underline{r}^{(0)}$$

$$= A^{-1}\left(\sum_{i=0}^{k}\alpha_i A^i\right)\underline{r}^{(0)} = \left(\sum_{i=0}^{k}\alpha_i A^{i-1}\right)\underline{r}^{(0)} = \left(\sum_{i=0}^{k}\alpha_i A^i\right)A^{-1}\underline{r}^{(0)}$$
$$= p_k(A)\,A^{-1}\underline{r}^{(0)} = p_k(A)\left(\underline{x} - \underline{x}^{(0)}\right). \tag{7.3}$$

Consequently, the error can be characterized with the help of $p_k(A)$. □

Remark 7.3. S.p.d. matrices and the CG method. Consider the case that A is symmetric and positive definite. Then, one gets from (7.3)

$$\left\|\underline{x} - \underline{x}^{(k)}\right\|_A = \left\|p_k(A)\left(\underline{x} - \underline{x}^{(0)}\right)\right\|_A.$$

The iterate $\underline{x}^{(k)}$ of the conjugate gradient method is the minimizer of $\left\|\underline{x} - \underline{x}^{(k)}\right\|_A$ in $\underline{x}^{(0)} + K_k\left(\underline{r}^{(0)}, A\right)$, see Theorem 6.12. Hence

$$\left\|\underline{x} - \underline{x}^{(k)}\right\|_A = \min_{p_k \in P_k, p_k(0)=1}\left\|p_k(A)\left(\underline{x} - \underline{x}^{(0)}\right)\right\|_A, \tag{7.4}$$

since $p_k(A)$ is the only parameter in the expression on the right-hand side. From

$$\begin{aligned}
&\left\|p_k(A)\left(\underline{x} - \underline{x}^{(0)}\right)\right\|_A \\
&= \left(\left(p_k(A)\left(\underline{x} - \underline{x}^{(0)}\right)\right)^T A\left(p_k(A)\left(\underline{x} - \underline{x}^{(0)}\right)\right)\right)^{1/2} \\
&= \left(\left(A^{1/2}p_k(A)\left(\underline{x} - \underline{x}^{(0)}\right)\right)^T\left(A^{1/2}p_k(A)\left(\underline{x} - \underline{x}^{(0)}\right)\right)\right)^{1/2} \\
&= \left(\left(p_k(A)A^{1/2}\left(\underline{x} - \underline{x}^{(0)}\right)\right)^T\left(p_k(A)A^{1/2}\left(\underline{x} - \underline{x}^{(0)}\right)\right)\right)^{1/2} \\
&= \left\|p_k(A)A^{1/2}\left(\underline{x} - \underline{x}^{(0)}\right)\right\|_2 \\
&\le \left\|p_k(A)\right\|_2\left\|A^{1/2}\left(\underline{x} - \underline{x}^{(0)}\right)\right\|_2 = \left\|p_k(A)\right\|_2\left\|\underline{x} - \underline{x}^{(0)}\right\|_A,
\end{aligned}$$

it follows by insertion in (7.4) that

$$\frac{\left\|\underline{x} - \underline{x}^{(k)}\right\|_A}{\left\|\underline{x} - \underline{x}^{(0)}\right\|_A} \le \min_{p_k \in P_k, p_k(0)=1}\left\|p_k(A)\right\|_2. \tag{7.5}$$

The right-hand side of (7.5) is the same as the right-hand side of (7.2). □

Lemma 7.4. Characterization of $\|p_k(A)\|_2$ for normal matrices. *If $A \in \mathbb{R}^{n\times n}$ is a normal matrix, see Definition 2.14, then*

$$\|p_k(A)\|_2 = \max_{\substack{\lambda \text{ is} \\ \text{eigenvalue of } A}} |p_k(\lambda)| .$$

Proof. Let $p_k \in P_k$ be an arbitrary polynomial with $p_k(0) = 1$. Then, one obtains by using Remark 2.15

$$\begin{aligned}
\|p_k(A)\|_2 &= \left\|p_k\left(Q^T \Lambda Q\right)\right\|_2 = \left\|\sum_{i=0}^{k} \alpha_i \left(Q^T \Lambda Q\right)^i\right\|_2 \\
&= \left\|Q^T\left(\sum_{i=0}^{k} \alpha_i \Lambda^i\right) Q\right\|_2 = \left\|Q^T p_k(\Lambda) Q\right\|_2 = \|p_k(\Lambda)\|_2 ,
\end{aligned}$$

since

$$\left(Q^T \Lambda Q\right)^i = \left(Q^T \Lambda \underbrace{Q\right)(Q^T}_{=I} \Lambda \underbrace{Q\right)\left(Q^T}_{=I} \Lambda Q\right) \ldots \left(Q^T \Lambda Q\right) = Q^T \Lambda^i Q$$

and the $\|\cdot\|_2$ norm is invariant with respect to the multiplication with unitary matrices. The matrix $p_k(\Lambda)$ is diagonal with the entries $p_k(\lambda_i)$. Hence

$$\|p_k(A)\|_2 = \max_{1 \le i \le n} |p_k(\lambda_i)|$$

by the definition of the spectral norm. ∎

Remark 7.5. Chebyshev polynomials. For proving the convergence theorem, Chebyshev[1] polynomials of first kind will be used, see also the lecture notes of Numerical Mathematics I,

$$\begin{aligned}
T_k(x) &= \cos(k \arccos(x)) \\
&= x^k - \binom{k}{2} x^{k-2}\left(1-x^2\right) + \binom{k}{4} x^{k-4}\left(1-x^2\right)^2 \\
&\quad - \binom{k}{6} x^{k-6}\left(1-x^2\right)^3 \ldots, \quad x \in [-1,1].
\end{aligned}$$

In particular, it is $T_k(x) \in [-1,1]$ for $x \in [-1,1]$,

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1.$$

The domain of definition of $T_k(x)$ can be extended to $|x| > 1$. It is, with $i = \sqrt{-1}$,

$$\arccos(x) = \frac{1}{i} \ln\left(x + \sqrt{x^2 - 1}\right), \quad x \in \mathbb{R},$$

such that

$$T_k(x) = \cos\left(\frac{k}{i} \ln\left(x + \sqrt{x^2 - 1}\right)\right). \tag{7.6}$$

For $x > 1$, one has

[1] Pafnuty Lvovich Chebyshev (1821 – 1894)

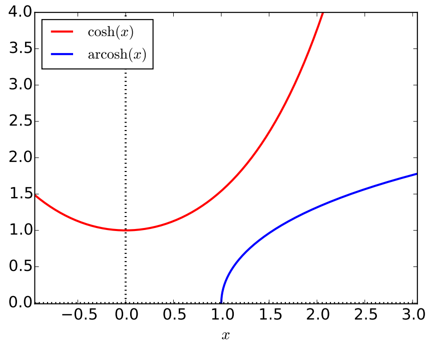

Fig. 7.1 Sketches of $\cosh(x)$ and $\text{arcosh}(x)$.

$$\ln\left(x+\sqrt{x^2-1}\right) = \text{arcosh}\,(x)\,, \tag{7.7}$$

see Figure 7.1 for sketches of the hyperbolic functions. Using that

$$e^{iz} = \cos(z) + i\sin(z) \quad \Longrightarrow \quad \cos(z) = \frac{e^{iz}+e^{-iz}}{2} = \cosh\,(iz) \quad \forall\, z \in \mathbb{C},$$

gives

$$\cos\left(\frac{z}{i}\right) = \cos\,(-iz) = \cosh\,(z)\,, \quad z \in \mathbb{C}.$$

With (7.6) and (7.7), it follows that

$$T_k\,(x) = \cosh\,(k\ \text{arcosh}\,(x)) \ \text{ for } x > 1.$$

For symmetry reasons, one obtains for $x < -1$

$$T_k\,(x) = (-1)^k \cosh\,(k\ \text{arcosh}\,(-x))\,. \tag{7.8}$$

□

Theorem 7.6. Estimate of the rate of convergence for the CG method. *Let A be symmetric and positive definite with $\lambda_{\min} < \lambda_{\max}$, then*

$$\min_{p_k \in P_k, p_k(0)=1} \|p_k(A)\|_2 \le 2\left(\frac{\sqrt{\kappa_2\,(A)}-1}{\sqrt{\kappa_2\,(A)}+1}\right)^k. \tag{7.9}$$

Consequently, it is for the CG method

$$\frac{\left\|\underline{x}-\underline{x}^{(k)}\right\|_A}{\left\|\underline{x}-\underline{x}^{(0)}\right\|_A} \le 2\left(\frac{\sqrt{\kappa_2\,(A)}-1}{\sqrt{\kappa_2\,(A)}+1}\right)^k. \tag{7.10}$$

Proof. The idea of the proof consists in constructing a special polynomial which gives the estimate since

$$\min_{p_k\in P_k, p_k(0)=1} \|p_k(A)\|_2 \le \left\|p_{k,\text{special}}(A)\right\|_2 .$$

Let $\lambda_{\min}$ be the smallest and $\lambda_{\max}$ be the largest eigenvalue of A. Consider the linear function

$$\lambda \ : \ \mathbb{R} \to \mathbb{R}, \quad t \mapsto \frac{\lambda_{\min}+\lambda_{\max}}{2} + \frac{\lambda_{\max}-\lambda_{\min}}{2} t.$$

In particular, the restriction $t \in [-1, 1]$ gives $\lambda \in [\lambda_{\min}, \lambda_{\max}]$. The root of $\lambda(t)$ is denoted by t_0. It is

$$t_0 = -\frac{\lambda_{\min}+\lambda_{\max}}{\lambda_{\max}-\lambda_{\min}} = -\frac{\kappa_2(A)+1}{\kappa_2(A)-1} < -1, \tag{7.11}$$

where one uses that for symmetric positive definite matrices that $\kappa_2(A) = \lambda_{\max}/\lambda_{\min}$. Denoting by $t(\lambda)$ the inverse function, one defines the special polynomial

$$p_k(\lambda) = \frac{T_k(t(\lambda))}{T_k(t(0))} =: \frac{T_k(t)}{T_k(t_0)} \in P_k.$$

Then, it is $p_k(0) = T_k(t_0)/T_k(t_0) = 1$. It is by Lemma 7.4 and since $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ for all eigenvalues of A (the maximum does not decrease if it is searched in a larger set)

$$\begin{aligned}\|p_k(A)\|_2 &= \max_{\lambda \text{ is eigenvalue of } A} |p_k(\lambda)| \le \max_{\lambda\in[\lambda_{\min},\lambda_{\max}]} |p_k(\lambda)| = \max_{t\in[-1,1]} \frac{|T_k(t)|}{|T_k(t_0)|} \\ &= \frac{1}{|T_k(t_0)|} \underbrace{\max_{t\in[-1,1]} |T_k(t)|}_{\le 1} \le \frac{1}{|T_k(t_0)|}.\end{aligned} \tag{7.12}$$

For estimating this term, consider (7.8) since $t_0 < -1$:

$$|T_k(t_0)| = \left|(-1)^k \cosh(k \underbrace{\operatorname{arcosh}(-t_0)}_{\omega_0})\right| = |\cosh(k\omega_0)| = \frac{e^{k\omega_0}+e^{-k\omega_0}}{2}.$$

One has to estimate this term from below. Consider first the case $k = 1$. Since $-t_0 > 1$, one has, by applying the inverse function,

$$\frac{e^{\omega_0}+e^{-\omega_0}}{2} = \cosh(\omega_0) = \cosh(\operatorname{arcosh}(-t_0)) = -t_0,$$

from what $e^{\omega_0} + e^{-\omega_0} = -2t_0$ follows. This equation is quadratic in e^{ω_0} with the solution

$$e^{\omega_0} = \underbrace{-t_0}_{>1} \pm \sqrt{t_0^2-1}.$$

For estimating $|T_k(t_0)|$, one obtains a sharper estimate if the larger one of these two values is considered, see (7.13) below. One gets with (7.11) and the binomial theorem

$$\begin{aligned} e^{\omega_0} &= -t_0 + \sqrt{t_0^2-1} = \frac{\kappa_2(A)+1}{\kappa_2(A)-1} + \sqrt{\frac{(\kappa_2(A)+1)^2-(\kappa_2(A)-1)^2}{(\kappa_2(A)-1)^2}} \\ &= \frac{\kappa_2(A)+2\sqrt{\kappa_2(A)}+1}{\kappa_2(A)-1} = \frac{\left(\sqrt{\kappa_2(A)}+1\right)^2}{\left(\sqrt{\kappa_2(A)}+1\right)\left(\sqrt{\kappa_2(A)}-1\right)} = \frac{\sqrt{\kappa_2(A)}+1}{\sqrt{\kappa_2(A)}-1}.\end{aligned}$$

Now, $|T_k(t_0)|$ is estimated from below

$$|T_k(t_0)| = \frac{e^{k\omega_0} + e^{-k\omega_0}}{2} > \frac{e^{k\omega_0}}{2} = \frac{\left(e^{\omega_0}\right)^k}{2} = \frac{1}{2}\left(\frac{\sqrt{\kappa_2(A)}+1}{\sqrt{\kappa_2(A)}-1}\right)^k. \tag{7.13}$$

Inserting this estimate in (7.12) finishes the proof of (7.9).

Estimate (7.10) is obtained by inserting (7.9) in (7.5). ■

Remark 7.7. The case $\lambda_{\min} = \lambda_{\max} = \lambda$. From Remark 2.15, it follows that for $\lambda_{\min} = \lambda_{\max} = \lambda$

$$A = Q^T \lambda I Q = \lambda Q^T Q = \lambda I,$$

i.e., A is a multiple of the identity matrix. In this case, the linear system of equations can be solved directly, without using the CG method. Choosing $p_1(x) = -x/\lambda + 1$, then it is $p_1(0) = 1$, $p_1(\lambda) = 0$ and consequently $\|p_1(A)\|_2 = 0$, see Lemma 7.4. That means, the CG method converges in one iteration. □

Remark 7.8. Connection of the number of iterations and the spectral condition number. To guarantee the reduction of the error by a factor $0 < \eta < 1$ on the basis of (7.5) and estimate (7.9) from Theorem 7.6, the condition

$$2\left(\frac{\sqrt{\kappa_2(A)}-1}{\sqrt{\kappa_2(A)}+1}\right)^k \le \eta$$

must be satisfied. The number of iterations to achieve this condition is

$$k \ge \frac{|\ln(\eta/2)|}{\left|\ln\left(\frac{\sqrt{\kappa_2(A)}-1}{\sqrt{\kappa_2(A)}+1}\right)\right|} = \frac{-\ln(\eta/2)}{\ln\left(\frac{\sqrt{\kappa_2(A)}+1}{\sqrt{\kappa_2(A)}-1}\right)}.$$

If $\kappa_2(A)$ is large, then a power series expansion of the logarithm

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots,$$

gives, using only the linear term,

$$\begin{aligned}
\ln\left(\frac{\sqrt{\kappa_2(A)}+1}{\sqrt{\kappa_2(A)}-1}\right) &= \ln\left(\frac{1+\frac{1}{\sqrt{\kappa_2(A)}}}{1-\frac{1}{\sqrt{\kappa_2(A)}}}\right) \\
&= \ln\left(1+\frac{1}{\sqrt{\kappa_2(A)}}\right) - \ln\left(1-\frac{1}{\sqrt{\kappa_2(A)}}\right) \\
&\approx \frac{2}{\sqrt{\kappa_2(A)}}.
\end{aligned}$$

That means, an approximation of the upper bound of the number of iterations to reduce the error by the factor η is

$$k \approx \frac{-\ln(\eta/2)}{2}\sqrt{\kappa_2(A)}.$$

The dependency on $\mathcal{O}\left(\sqrt{\kappa_2(A)}\right)$ can be observed in fact sometimes. However, often the convergence of the CG method is considerably faster than predicted by the upper bound (7.10). One can derive better and sharper bounds if the distribution of all eigenvalues is considered instead only the smallest and the largest one, which are needed for computing the spectral condition number. □

Remark 7.9. Round-off errors. For studying the behavior of the CG method in practice, one has to take into account in the analysis the round-off errors that are committed due to the finite precision arithmetic. The accumulation of round-off errors might lead to an increasing loss of the property of the computed vectors to be A-conjugate. Then, the computed solution might be only a quite inaccurate approximation of the solution of (1.1). □