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 if one can get information about the iterate $\mathbf{x}^{(k)}$ for $k < n$. □

Remark 7.2 *Starting point of the convergence analysis.* The basis of the convergence analysis for Krylov subspace methods is the following observation: $\mathbf{z} \in K_k\left(\mathbf{r}^{(0)}, A\right)$ is equivalent to $\mathbf{z} = q_{k-1}(A)\,\mathbf{r}^{(0)}$, 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

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

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

Considering the methods which are based on the minimzation of the residual, see Chapter 5, one has now

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

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

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

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

$$\begin{aligned}\mathbf{x} - \mathbf{x}^{(k)} &= A^{-1}\mathbf{b} - A^{-1}\left(\mathbf{b} - \mathbf{r}^{(k)}\right) = A^{-1}\mathbf{r}^{(k)} = A^{-1}p_k(A)\,\mathbf{r}^{(0)} \\ &= A^{-1}\left(\sum_{i=0}^{k}\alpha_i A^i\right)\mathbf{r}^{(0)} = \left(\sum_{i=0}^{k}\alpha_i A^{i-1}\right)\mathbf{r}^{(0)} = \left(\sum_{i=0}^{k}\alpha_i A^i\right)A^{-1}\mathbf{r}^{(0)} \\ &= p_k(A)\,A^{-1}\mathbf{r}^{(0)} = p_k(A)\left(\mathbf{x} - \mathbf{x}^{(0)}\right). \qquad (7.3)\end{aligned}$$

□

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)

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

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

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

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

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

it follows that

$$\frac{\left\| \mathbf{x} - \mathbf{x}^{(k)} \right\|_A}{\left\| \mathbf{x} - \mathbf{x}^{(0)} \right\|_A} \le \min_{p_k \in P_k, p_k(0)=1} \left\| p_k(A) \right\|_2 . \tag{7.4}$$

The right hand side of (7.4) 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*

$$\min_{p_k \in P_k, p_k(0)=1} \left\| p_k(A) \right\|_2 = \min_{p_k \in P_k, p_k(0)=1} \max_{\lambda \text{ is eigenvalue of } A} \left| p_k(\lambda) \right| .$$

Proof: Let $p_k \in P_k$ be an arbitrary polynomial with $p_k(0) = 1$. Then

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

since

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

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

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

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 use, 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 $T_k \in [-1,1]$ 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

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

such that

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

For $x > 1$, one has

$$\ln\left(x + \sqrt{x^2-1}\right) = \operatorname{arcosh}(x)$$

and from

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

it follows that

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

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

$$T_k(x) = (-1)^k \cosh(k \operatorname{arcosh}(-x)). \tag{7.5}$$

□

Theorem 7.6 Estimate of the rate of convergence for s.p.d. matrices. *Let A be symmetric and positive definit. Then*

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

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

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

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

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

[1]Pafnuty Lvovich Chebyshev (1821 – 1894)

where one uses that for symmetric positive definite matrices $\kappa_2(A) = \lambda_{\max}/\lambda_{\min}$. 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 $p_k(0) = T_k(t_0)/T_k(t_0) = 1$. 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)

$$\begin{aligned} \|p_k(A)\|_2 &= \max_{\lambda \text{ is eigenvalue of } A} |p_k(\lambda)| \le \max_{\lambda\in[\lambda_{\min},\lambda_{\max}]} |p_k(\lambda)| = \max_{t\in[-1,1]} \frac{|T_k(t)|}{|T_k(t_0)|} \\ &= \frac{1}{|T_k(t_0)|} \underbrace{\max_{t\in[-1,1]} |T_k(t)|}_{\le 1} \le \frac{1}{|T_k(t_0)|}. \end{aligned} \tag{7.6}$$

For estimating this term, consider (7.5) 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. Since $-t_0 > 1$, one has

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

from what $e^{\omega_0} + e^{-\omega_0} = -2t_0$ follows. This is a quadratic equation 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, i.e.,

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

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

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

Inserting this estimate into (7.6) finishes the proof. ■

Remark 7.7 *Connection of the number of iterations and the spectral condition number.* To guarantee the reduction of the error by a factor $\eta < 1$, using the estimate from Theorem 7.6,

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

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 gives

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

That means, the expected number of iterations to reduce the error by the factor η increases with

$$k \approx \frac{-\ln(\eta/2)}{2}\sqrt{\kappa_2(A)} = \mathcal{O}\left(\sqrt{\kappa_2(A)}\right).$$

This bevavior can be observed in fact in many situations, e.g., for linear systems of equations arising in discretizations of partial differential equations. □