AN ERROR ESTIMATE FOR THE APPROXIMATION OF $\cosh(\sqrt{A})v$ IN THE KRYLOV SUBSPACE *

V. SIMONCINI† AND L. ELDÉN‡

Abstract. Krylov subspaces are a valuable tool for approximating the action of a matrix function to a vector, when the matrix has large dimension. In this note we provide an error estimate for the approximation of the function $\cosh(\sqrt{\lambda})$ by generalizing a result obtained for the exponential function.

1. Introduction. Matrix functions in the form $f(L)$, where L is a square matrix and f is a function such that $f(L)$ is well defined, play an increasingly important role in the approximation of the solution to evolution partial differential equations and in certain control problems, see, e.g., [7, 9, 5] and references therein. In most cases where the matrix L has large dimension n, the operation $f(L)v$ is actually required, where v is a vector. In this setting, approximations of $f(L)v$ by subspace projection turn out to be particularly appealing. Here we focus on the use of the Krylov subspace $K_m(L, v) = \text{span}\{v, Lv, \ldots, L^{m-1}v\}$ of size m, which has proven to be particularly well suited for the approximation of functions such the exponential; see, e.g., [3, 6, 8, 10, 11]. In the following we assume that L is symmetric and positive definite. Let V_m be a matrix whose orthonormal columns span $K_m(L, v)$, with $v = V_m e_1\|v\|$, and let $T_m = V_m^T L V_m$. The approximation in the Krylov subspace can be written as

$$V_m f(T_m) e_1 \|v\| \;\approx\; f(A)v,$$

where for $m \ll n$, the matrix function $f(T_m)$ can be computed efficiently by means of an eigendecomposition. In this note we are concerned with estimates for the error norm $\|f(A)v - V_m f(T_m)e_1\|v\|\|$ when $f(\lambda) = \cosh(\sqrt{\lambda})$, as a function of the subspace dimension m. The approximation procedure will be effective if a sufficiently accurate approximation is obtained for m much smaller than n. In a worst case scenario, we show that such an effectiveness fully depends on the spectral interval of the matrix L. In practice, convergence can be faster, but this depends on the eigenvalue distribution within the spectral interval, and on the decomposition of v in terms of eigenvectors of L. Without more detailed spectral information on the problem, the worst case scenario may be considered the best possible estimate.

2. The error norm estimate. It was shown in [2] that the error norm for approximating the exponential function in the Krylov subspace can be effectively estimated. In this note we extend this result to the case of the hyperbolic cosine of the scaled square root, defined as $\cosh(z_1\sqrt{\lambda}) = \frac{1}{2}(\exp(z_1\sqrt{\lambda}) + \exp(-z_1\sqrt{\lambda}))$, where z_1 is some positive constant. This result was used in [4].

PROPOSITION 2.1. *Let L be symmetric and positive definite, and $\lambda_{\min}$, $\lambda_{\max}$ be the extreme eigenvalues of L. For $z_1 > 0$, let $a = z_1\sqrt{\lambda_{\max}}$. Let V_m be such that range$(V_m) = K_m(L, v)$, $T_m = V_m^T L V_m$ and $\beta_0 = \|v\|$. If $m \le a/2$, then for large a, the error $\mathcal{E}_m = \cosh(z_1\sqrt{L})v - V_m \cosh(z_1\sqrt{T_m})e_1\beta_0$ satisfies*

$$\|\mathcal{E}_m\| \approx \frac{a}{2m} \exp(-a((2m)^2/(2a^2) + O((2m/a)^4)),$$

*Version of 14 Jan 2010.

†Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato 5, I-40127 Bologna, Italy and CIRSA, Ravenna, Italy (valeria@dm.unibo.it).

‡Department of Mathematics, Linköping University, Sweden. (laeld@math.liu.se)

where $\approx$ means that higher order terms are neglected.

Proof. The proof is based on similar results by Druskin & Knizhnerman [2]. For $z_1 > 0$ and $\tau = 2/\sqrt{\lambda_{\max}}$, let $B = I - \frac{\tau^2}{2}L$, and $a := 2z_1/\tau = z_1\sqrt{\lambda_{\max}}$. We first show that $\cosh(z_1\sqrt{L})$ can be expanded in Chebychev series as

$$\cosh(z_1\sqrt{L}) = I_0(a) + 2\sum_{k=1}^{\infty}(-1)^k I_{2k}(a)\mathcal{T}_k(B),$$

where $\mathcal{T}_k$ is the Chebyshev polynomial of degree k, and I_j is the modified Bessel function. With some abuse of notation, let θ be such that $B = \cos(2\theta)$. We recall the following expansion (see [1, n. 9.6.35, p. 376]):

$$e^{a\sin\theta} = I_0(a) + 2\sum_{k=0}^{\infty}(-1)^k I_{2k+1}(a)\sin\left((2k+1)\theta\right) + 2\sum_{k=1}^{\infty}(-1)^k I_{2k}(a)\cos(2k\theta).$$

Moreover, $I_k(a) = (-1)^k I_k(-a)$ for $k = 0, 1, \ldots$. Then we can write

$$\begin{aligned}
\cosh(a\sin\theta) &= \frac{1}{2}(e^{a\sin\theta} + e^{-a\sin\theta}) \\
&= \frac{1}{2}(I_0(a) + I_0(-a)) \\
&\quad + \sum_{k=0}^{\infty}(-1)^k I_{2k+1}(a)\sin(2k+1)\theta + \sum_{k=1}^{\infty}(-1)^k I_{2k}(a)\cos(2k\theta) \\
&\quad + \sum_{k=0}^{\infty}(-1)^k I_{2k+1}(-a)\sin(2k+1)\theta + \sum_{k=1}^{\infty}(-1)^k I_{2k}(-a)\cos(2k\theta) \\
&= I_0(a) + 2\sum_{k=1}^{\infty}(-1)^k I_{2k}(a)\cos(2k\theta).
\end{aligned}$$

Since $\theta = \frac{1}{2}\text{arcos}(B)$, it follows that $\cos(2k\theta) = \mathcal{T}_k(B)$. Moreover, $\cosh(a\sin\theta) = \cosh(a\sqrt{1-\cos^2\theta}) = \cosh(a\sqrt{(I-B)/2})$. Using the fact that $z_1\sqrt{L} = a\sqrt{(I-B)/2}$ the expansion follows.

We next prove the new estimate. Since $\|B\| \le 1$, we have $\|\mathcal{T}_k(B)\| \le 1$. Therefore

$$\begin{aligned}
\|\cosh(z_1\sqrt{L})v - V_m\cosh(z_1\sqrt{T_m})e_1\beta_0\| &= \|2\sum_{k=m}^{\infty}(-1)^k I_{2k}(a)\mathcal{T}_k(B)\| \\
&\le 2\sum_{k=m}^{\infty}|I_{2k}(a)|.
\end{aligned}$$

It was shown in Druskin & Knizhnerman ([2, proof of Th. 4]) that for $2m/a \le 1$,

$$\begin{aligned}
\sum_{k=m}^{\infty}|I_{2k}(a)| &\le c(a)\sum_{k=m}^{\infty}\exp(((2k)^2+a^2)^{\frac{1}{2}} - 2k\text{arsh}(2k/a)) \\
&\le c_1(a)\frac{a}{2m}\exp(-a((2m)^2/(2a^2) + O((2m/a)^4)),
\end{aligned}$$

from which the result follows. □

The proposition shows that convergence indeed depends on the norm of $L^{1/2}$, that is on the upper extreme of its spectral interval.

REFERENCES

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, 1965.

[2] V. Druskin and L. Knizhnerman. Two polynomial methods of calculating functions of symmetric matrices. *U.S.S.R. Comput. Math. Math. Phys.*, 29:112–121, 1989.

[3] V. Druskin and L. Knizhnerman. Krylov subspace approximation of eigenpairs and matrix functions in exact and computer arithmetic. *Numerical Linear Algebra with Appl.*, 2(3):205–217, 1995.

[4] L. Eldèn and V. Simoncini. Solving ill-posed Cauchy problems by a Krylov subspace method. *Inverse Problems*, 25(6):–, 2009.

[5] Andreas Frommer and Valeria Simoncini. Matrix functions. In Wil H. A. Schilders, Henk A. van der Vorst, and Joost Rommes, editors, *Model Order Reduction: Theory, Research Aspects and Applications*, Mathematics in Industry. Springer, Heidelberg, 2008.

[6] E. Gallopoulos and Y. Saad. Efficient solution of parabolic equations by Krylov approximation methods. *SIAM J. Sci. Stat. Comput.*, 13(5):1236–1264, 1992.

[7] Nicholas J. Higham. *Matrix Functions – Theory and Applications*. SIAM, Philadelphia, USA, 2008.

[8] M. Hochbruck and C. Lubich. On Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 34(5):1911–1925, 1997.

[9] Marlis Hochbruck and Alexander Ostermann. Exponential integrators. *Acta Numerica*, pages –, 2010. To appear.

[10] Y. Saad. Analysis of some Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 29:209–228, 1992.

[11] R. B. Sidje. Expokit: A Software Package for Computing Matrix Exponentials. *ACM Transactions on Math. Software*, 24(1):130–156, 1998.