Differentiating the Matrix Exponential Function

Michael Eiermann
Institut für Numerische Mathematik und Optimierung
Technische Universität Bergakademie Freiberg, Germany

Let $A = A(\boldsymbol{c})$ be an $n \times n$ matrix depending on a parameter vector $\boldsymbol{c}$ and let $\boldsymbol{b}$ be an n-dimensional vector. New algorithms are developped for computing the matrix-vector products $\boldsymbol{v} \mapsto J\boldsymbol{v}$ and $\boldsymbol{w} \mapsto J^T\boldsymbol{w}$, where J denotes the derivative of $u(t, \boldsymbol{c}) = \exp(tA(\boldsymbol{c}))\boldsymbol{b}$ with respect to $\boldsymbol{c}$. Problems of this type arise in the context of parameter identification problems. Here, A is usually very large and sparse, i.e., none of the large, but dense matrices, $\exp(tA)$ or J, can be computed explicitly or stored. The algorithms are based on the approximation of the matrix exponential by partial sums of its Taylor or Chebyshev series and require only matrix-vector multiplications by A. To avoid numerical cancellation, a scaling strategy is included.

We analyze the truncation errors of these algorithms and describe their behavior in finite precision arithmetic. We discuss the selection of the two algorithmic parameters, namely the truncation index of the series and the amount of scaling.

We finally illustrate the use of these algorithms by a parameter identification problem for the heat equation and compare their computational efficiency to those of other approaches which are based on rational approximations of the exponential function, the solution of ordinary differential equations, the projection onto Krylov subspaces, or the popular adjoint method.