Discretised dynamical low-rank approximation in the presence of small singular values

Emil Kieri

(joint work with Christian Lubich, Hanna Walach)

We prove error estimates for a novel time-stepping scheme for low-rank matrix and tensor differential equations. The estimate is robust with respect to small singular values. When a singular value in the approximation approaches zero, standard time-stepping schemes break down. We show that the new method solves this problem.

Low-rank approximations have had much success in the field of quantum dynamics, in particular through the multi-configurational time-dependent Hartree (MCTDH) method [5]. However, also time-stepping schemes for MCTDH have difficulties in the presence of small singular values. MCTDH uses the Tucker format to construct low-rank approximations of tensors. In this work we use a different low-rank tensor format, known as tensor trains or matrix product states.

We consider the low-rank approximation of a large, time-dependent tensor $A(t) \in \mathbb{C}^{n_1\times\cdots\times n_d}$, given via a tensor differential equation

$$\dot{A}(t) = F(t, A(t)), \qquad A(0) = A_0 \in \mathbb{C}^{n_1\times\cdots\times n_d}.$$

If we can approximate $A(t)$ by a rank-r tensor train, the amount of data required to represent $A(t)$ would be reduced from $\mathcal{O}(n^d)$ to $\mathcal{O}(dr^2n)$, with $n = \max n_i$. To keep the notation simple we will in this note only consider the matrix case, i.e., $d = 2$, and aim at approximating $A(t)$ by a rank-r matrix. The results extend to low-rank tensors in the tensor train format with arbitrary d.

Commonly, the singular values of a matrix decay without a distinct gap. This means that the last included and first neglected singular values, σ_r and σ_{r+1}, are of similar size. σ_{r+1} represents neglected information, and if it is not small the low-rank approximation will introduce a large error. We should therefore expect also σ_r to be small. In this work we prove that the splitting scheme is robust in this situation: *If the exact solution is an ε-perturbation of a rank-r matrix, the error can be bounded in terms of ε and the time step, independently of the smallness of σ_r.* For a more precise statement and a proof of this result, see [1].

We approximate $A(t)$ by a matrix of rank r using the SVD-like decomposition

$$A(t) \approx Y(t) = U(t)S(t)V(t)^*,$$

where $U \in \mathbb{C}^{n_1\times r}$ and $V \in \mathbb{C}^{n_2\times r}$ have orthonormal columns and $S \in \mathbb{C}^{r\times r}$. We denote the manifold of rank-r matrices by $\mathcal{M}_r$ and its tangent space at Y by $\mathcal{T}_Y\mathcal{M}_r$. We then determine the time-evolution of $Y(t)$ using the Dirac–Frenkel time-dependent variational principle,

$$\dot{Y}(t) = P(Y(t))F(t, Y(t)), \qquad Y(0) = Y_0, \tag{1}$$

where $P(Y)$ is the orthogonal projection onto the tangent space $\mathcal{T}_Y\mathcal{M}_r$. This can also be seen as a Galerkin condition on the tangent space. Subject to a gauge condition, (1) determines a system of differential equations for the matrices U, S and V [2],

$$\dot{U}(t) = (I - U(t)U(t)^*)F(t, Y(t))V(t)S(t)^{-1},$$
$$\dot{S}(t) = U(t)^*F(t, Y(t))V(t),$$
$$\dot{V}(t) = (I - V(t)V(t)^*)F(t, Y(t))^*U(t)S(t)^{-*}.$$

We note that this system is stiff if σ_r is small, and does not have a well-defined solution in the limit $\sigma_r \to 0$.

The projection onto the tangent space can be decomposed as

$$P(Y)Z = ZVV^* - UU^*ZVV^* + UU^*Z, \qquad Y = USV^*, \qquad Z \in \mathbb{C}^{n_1 \times n_2}.$$

Recently, a time-stepping scheme based on this splitting was proposed [3]. A similar scheme for the tensor train case has also been constructed [4]. Error bounds in terms of the time step h are available by standard theory for splitting methods, but unfortunately these estimates break down when $\sigma_r \to 0$. Such a break-down is, however, not observed in numerical experiments. The splitting scheme possesses a remarkable exactness property, which gives a first theoretical indication of its robustness: If $A(t) \in \mathcal{M}_r$ for all t and its time-derivative $\dot{A}(t) = F(t)$ is given independently of $A(t)$, then the splitting method is exact for any h and independently of σ_r. Our analysis unifies this property with the standard error estimates.

The error estimate requires F to be Lipschitz continuous. This is a considerable limitation in a quantum dynamics context, since for a discretisation of the Schrödinger equation with spatial step size Δx the Lipschitz constant will be of order Δx^{-2}. This suggests that very small time steps would be needed. Such a time step restriction is however not observed in numerical experiments. The method seems to be robust for partial differential equations, and it would be of interest to extend the theory also to this situation.

References

[1] E. Kieri, Ch. Lubich, and H. Walach, *Discretised dynamical low-rank approximation in the presence of small singular values*. Preprint (2015).

[2] O. Koch and Ch. Lubich, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl. **29** (2007), 434–454.

[3] Ch. Lubich and I. V. Oseledets, *A projector-splitting integrator for dynamical low-rank approximation*, BIT **54** (2014), 171–188.

[4] Ch. Lubich, I. V. Oseledets, and B. Vandereycken, *Time integration of tensor trains*, SIAM J. Numer. Anal. **53** (2015), 917–941.

[5] H.-D. Meyer, U. Manthe, and L. S. Cederbaum, *The multi-configurational time-dependent Hartree approach*, Chem. Phys. Lett. **165** (1990), 73–78.