Matrix Equations and Tensor Techniques
IX
Perugia, September 9-10, 2021

Mixed precision recursive block diagonalization for bivariate functions of matrices

Stefano Massei *Leonardo Robol*

Department of Mathematics, University of Pisa, Italy

Matrix functions, denoted by $f(A)$, appear naturally in several applications, such as the analysis of complex networks and control theory. The efficient evaluation of matrix functions has been subject of a thourough study in recent years [1].

This concept extends extends quite naturally to the bivariate setting. Given two square matrices $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$ and a complex-valued function $f(x, y)$, the *bivariate matrix function* $f\{A, B^T\}$ [3] is a linear operator on $\mathbb{C}^{m\times n}$. As in the univariate case, the definition of $f\{A, B^T\}$ can be given, equivalently, in terms of (bivariate) Hermite interpolation, power series expansion and contour integration. In the latter formulation, let Λ_A and Λ_B be the spectra of A and B, respectively, and let $f(x, y)$ be analytic in an open neighborhood of $\Lambda_A \times \Lambda_B$; then, $f\{A, B^T\}$ is defined as

$$\begin{aligned} f\{A, B^T\} : \mathbb{C}^{m\times n} &\longrightarrow \mathbb{C}^{m\times n} \\ C &\longrightarrow f\{A, B^T\}(C) := \oint_{\Gamma_A} \oint_{\Gamma_B} f(x, y)(xI - A)^{-1} C (yI - B)^{-1}\, dxdy, \end{aligned}$$

with Γ_A, Γ_B closed contours enclosing Λ_A and Λ_B, respectively.

Several matrix equations and related numerical problems can be expressed as the evaluation of bivariate matrix functions: solving a Sylvester equations corresponds to evaluating $1/(x + y)$; computing the Frechét derivative of $f(z)$ at A can be rephrased as evaluating the divided difference at the matrices A and A^T, applied to the desired direction. Similarly, any matrix function of $I \otimes A + B^T \otimes I$ can be recast as a bivariate matrix function of the form $f(x, y) = h(x + y)$.

For univariate matrix functions, the Schur-Parlett algorithm provides a robust tool for evaluating $f(A)$ for a generic function. The approach does not extend easily to the bivariate setting.

We propose a numerically reliable method for computing $f\{A, B^T\}(C)$ for a general function $f(x, y)$ without requiring that A and/or B can be diagonalized with a well conditioned similarity transformation. In complete analogy to the univariate scenario, our procedure computes the Schur decompositions $A = Q_A T_A Q_A^*$ and $B = Q_B T_B Q_B^*$, so that the task boils down to evaluate the bivariate function for triangular coefficients:

$$f\{A, B^T\}(C) = Q_A f\{T_A, T_B^T\}(\widetilde{C}) Q_B^*, \qquad \widetilde{C} := Q_A^* C Q_B.$$

A generalized block recurrence is applied to retrieve $f\{T_A, T_B^T\}(\widetilde{C})$; the recursion requires to compute f on pairs of diagonal blocks of T_A and T_B^T and to solve Sylvester equations involving either diagonal blocks of T_A or of T_B. In view of the latter operation, we need to reorder the Schur forms of A and B such that distinct diagonal blocks have sufficiently separated eigenvalues. Finally, we evaluate f on the smallest diagonal blocks of T_A and T_B^T, the so-called *atomic blocks*, via a truncated bivariate Taylor expansion or, in the spirit of [2], with a randomized approximate diagonalization technique combined with high precision arithmetic. The procedure can be interpreted as an implicit (recursive) block-diagonalization strategy, where the eigenvectors matrices are not formed explicitly.

References

[1] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, 2008, .

[2] N. J. Higham and X. Liu, *A multiprecision derivative-free Schur–Parlett algorithm for computing matrix functions*, (2020).

[3] D. Kressner, *Bivariate matrix functions*, Operators and matrices, 8 (2014), pp. 449–466, .