Towards understanding Krylov subspace methods through examples

Erin Carson, Jörg Liesen, and Zdeněk Strakoš

Abstract

Taking the 1952 landmark paper of Hestenes and Stiefel on the conjugate gradient method (CG) as their historical starting point, Krylov subspace methods for solving linear algebraic systems have been around for more than 70 years. Tens of thousands of research articles on Krylov subspace methods and their applications have been published by authors coming from the most diverse scientific backgrounds. Nevertheless, many questions about the behavior of Krylov subspace methods both in exact arithmetic and finite precision computations remain open.

Based on the recent paper [1], which focusses on CG and the generalized minimal residual method (GMRES), this talk will discuss mathematically important as well as practically relevant phenomena in the context of Krylov subspace methods. Rather than giving yet another technical and algorithmic overview or taxonomy of the methods, we will present a sequence of computed examples that are organized around what we consider *the main points for the understanding* of Krylov subspace methods.

Both CG and GMRES, and Krylov subspace methods in general, are characterized by a nontrivial nonlinear behavior, which is their main mathematical asset as well as their beauty. This nonlinear behavior with respect to the given matrix and initial vector comes from using projections onto Krylov subspaces, which is equivalent to enforcing some form of optimality of the approximate solution, and which requires adaptation of the method to the given data (matrix and initial vector) at each iteration step. Adaptation to the hidden inner structure of the problem to be solved can lead to a significant speedup of the convergence in comparison with linear iterative methods that do not adapt to the problem. Moreover, nonlinear adaptation to the problem can not be captured by linear convergence bounds. The mathematical and practical consequences of this main point about Krylov subspace methods will appear in several of the examples in this talk.

In particular, it will be demonstrated and explained that CG adapts in a significant (and nonlinear) way to the eigenvalue distribution of the given matrix, how eigenvalue clusters and their positions affect the convergence of CG, that a smaller condition number not necessarily leads to faster convergence of CG, and how outlying eigenvalues can cause delay of convergence in finite precision computations. For the GMRES method we will discuss similar topics, including the role of the eigenvalue distribution and the influence of the initial residual on the convergence. But here we will see that the mathematical understanding is less developed than for CG.

Our main goal is to argue that, despite their long history and widespread use in practical applications, Krylov subspace methods should still be seen as not yet fully understood mathematical objects that are worth studying. Any progress in their understanding, even of their mathematical fundamentals, will bring us a step further in exploiting their full nonlinear computational potential.

References

[1] E. Carson, J. Liesen and Z. Strakoš, *Towards understanding CG and GMRES through examples*, Linear Algebra and its Applications, 692 (2024), pp. 241-291.