Randomized small-block Lanczos for null space computations

Daniel Kressner and Nian Shao

Abstract

Computing the null space of a large matrix A, having no particular structure beyond being sparse, is a challenging computational problem, especially if the nullity N—the dimension of the null space—is large.

When N = 1, standard choices include the Lanczos method for computing the smallest eigenvalue and eigenvector of $A^{\mathsf{T}}A$. When N > 1, the Lanczos method (with a single starting vector) becomes unreliable and prone to miss components of the null space. In fact, in exact arithmetic only a one-dimensional subspace of the null space can be extracted via Lanczos. Block Lanczos methods address this issue by utilizing a block of d > 1 starting vectors instead of a single starting vector. Common wisdom and existing convergence analyses all suggest to choose d not smaller than the size of the eigenvalue cluster of interest. Applied to null space computation this would require $d \ge N$ and in exact arithmetic this condition is indeed necessary. The presence of round-off error blurs the situation. As an example, consider the 420×420 diagonal matrix

$$A = \text{diag}(0, \dots, 0, 1, 2, \dots, 399), \tag{1}$$

that is, the first 21 diagonal entries are zero. When applying block Lanczos to A in order to compute its null space and choosing a block of d (Gaussian) random starting vectors, one would expect to obtain no more than d (approximate) zero eigenvalues and corresponding eigenvectors. However, when executing block Lanczos in double precision arithmetic without breakdown and declaring eigenvalues less than 10^{-4} as zero, we obtain the results reported in Table 1.

Table 1: Null space dimension obtained when applying block Lanczos with d random starting vectors to the matrix from (1).

Block size d	1	2	3	4	5	6	7	10	12	14	15	20	21
Dimension	11	11	16	12	15	15	21	20	21	21	21	21	21

While the null space dimension estimated by block Lanczos for small values of d is erratic and smaller than 21, it is always larger than d. In fact, already for d = 12, the correct null space dimension is detected. Experiments with other matrices lead to similar findings, suggesting that round-off error breaks some of the curse incurred by eigenvalue clusters, but it does not fully address the convergence issues of single-vector or small-block Lanczos methods either.

The situation described above is reminiscent of recent work [MMM2024] on the single-vector Lanczos method for low-rank approximation. Taking the randomness of the starting vector into account they establish a convergence result that still requires the (large) singular values to be distinct but the dependence of the complexity on the gaps between singular values is very mild, in fact logarithmic. Although single-vector or small-block Krylov subspace methods do not satisfy gap-independent convergence bounds, small random perturbations of A can easily break the adverse role of repeated singular values. As we have seen, perturbations due to round-off are not sufficient to achieve this effect, but (slightly larger) random perturbations will do, thanks to eigenvalue repulsion results.

In this work [KS2024], we propose a randomized small-block Lanczos method for null space calculation, sketched as Algorithm 1. Compared to the task of low-rank approximation, there are two major differences. First, in low-rank approximation, the largest few singular values are usually unknown and, in most applications, they do not form tight large clusters. In the context of null spaces, the desired singular values form one large cluster of zeros. Second, while convergence to the relevant invariant subspace is not necessary for low-rank approximation, such convergence is imperative to obtain a good null space approximation.

Algorithm 1: Randomized small-block Lanczos for null space computations
Input: Matrix A ∈ ℝ^{m×n} with m ≥ n and N zero singular values. Block size d. Perturbation parameter ε > 0. Number of iterations ℓ.
Output: Orthonormal basis V approximating null space of A.
Set B = A^TA + εD, with a diagonal matrix D having diagonal entries uniformly i.i.d. in [0, 1]; Draw a Gaussian random matrix Ω ∈ ℝ^{n×d};
Perform block Lanczos to compute the decomposition BZ_ℓ = Z_ℓT_ℓ + Q_{ℓ+1}E_{ℓ+1}, where Z_ℓ is an orthonormal basis of Krylov subspace span{Ω, BΩ, ..., B^{ℓ-1}Ω} and T_ℓ is block tridiagonal;
Compute an orthonormal basis V_Z for the invariant space of the N smallest eigenvalues of T_ℓ.

In the theoretical part, we examine the impact of the random perturbation ϵD . In particular, we present a new eigenvalue repulsion result for the perturbed zero eigenvalues of $A^{\mathsf{T}}A$. At the same time, the perturbation incurs a limit on the attainable accuracy of the null space approximation, and we quantify this effect by a perturbation analysis. For the single-vector case (d = 1), we achieve more refined results by analyzing the convergence of individual Ritz vectors and establish sharp bounds on the accuracy of the null space approximation.

In the numerical part, we incorporate several practical improvements, including preconditioning, restarting, and partial reorthogonalization, and provide various numerical results for applications such as the computation of connected graph components and cohomology. These numerical experiments not only validate our theoretical findings regarding the randomized single-vector Lanczos method but also highlight the efficiency of the small-block Lanczos method. Notably, it highlights when our method is preferable compared to other null space solvers, particularly when the nullity is substantial and memory constraints are a limiting factor.

References

- [KS2024] Daniel Kressner and Nian Shao. "A randomized small-block Lanczos method for largescale null space computations." arXiv preprint: 2407.04634.
- [MMM2024] Meyer, Raphael, Cameron Musco, and Christopher Musco. "On the unreasonable effectiveness of single vector krylov methods for low-rank approximation." Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA).