Inner-product free Krylov subspace methods for inverse problems

Malena Sabaté Landman, Ariana Brown, Julianne Chung, James G. Nagy

Abstract

We consider linear discrete inverse problems, which involve the reconstruction of hidden objects from possibly noisy indirect measurements. Rapid advances in technology and computation have resulted in enormous and growing data-sets, creating an urgent need for more efficient algorithms that can handle the increasing dimension of these problems while maintaining both reliability and, crucially, speed. New (and not so new) promising directions include reducing the working floating point arithmetic and increase parallelization, both of which can suffer from problems related to the use of inner-products in the algorithms. In this work, I present a **general class of Krylov methods** that are inherently **inner-product free**, while maintaining regularizing properties, making them a powerful and efficient alternative to traditional Krylov solvers in the context of **large-scale linear inverse problems**. Important applications appear, for example, in medical imaging (computed tomography), non-destructive testing of engineering designs, and geophysics (seismic exploration).

We write a linear system with additive noise as

$$Ax + e = b, \quad A \in \mathbb{R}^{m \times n},\tag{1}$$

where we typically consider e to be a realization of a white Gaussian random variable, and we focus on problems that are ill-posed in the sense that the system matrix A has decaying singular values which cluster at zero, but where A has an ill-determined numerical rank. Therefore, the reconstructions of x are typically very sensitive to perturbations in the measurements (a.k.a. noise) and need regularization.

Krylov subspace methods are a class of very popular projection methods to solve (1) which show very fast convergence and have inherent regularization properties when equipped with early stopping of the iterations [1]. In other words, the approximate solutions approach the true solution in the first iterations but, if the algorithms are not stopped, they continue to converge towards to unwanted solution of the least-squares problems associated with the noisy right hand side, which suffers badly from noise amplification: this is also referred to as semi-convergence. These are typically based on the stable construction of bases for Krylov subspaces, or search spaces for the solution x, defined as:

$$\mathcal{K}_k(C,d) = \mathcal{R}(V_k), \quad V_k = [d, Cd, \dots, C^{k-1}d], \tag{2}$$

which are related to the original linear system and where $\mathcal{R}(\cdot)$ represents the range of a matrix. Moreover, these are iterative methods, so that a new direction is added to the space $\mathcal{K}_k(C, d)$ at each iteration k. Usually, this is done using either the Arnoldi or Golub-Kahan bidiagonalization (GKB) algorithms, which, applied to (1), differ on the choice of the matrix-vector pair $\{C, d\}$. In particular, Arnoldi considers $\{A, b\}$; and GKB constructs basis for two Krylov subspaces; one taking $\{A^T A, A^T b\}$ and one taking $\{AA^T, b\}$. Both methods rely on the implicit construction of a QR factorization of the matrix V_k in (2), and they require the orthogonalization and normalization of the new basis vectors against the previous vectors in the basis, see, e.g. [2, Chapter 4]. However, there are scenarios where the inner-products required in this process can hinder the usability of the solvers. For example, in low precision arithmetic, standard Krylov solvers might break-down too early due to the norm of the new vector after orthogonalization being numerically zero in the given working precision, or over/under-flows can occur during norm computations. Moreover, inner-products can be a limiting factor for high performance computing, since they require global communication [3]. On the other hand, the most used inner-product free solvers, e.g. Chebyshev semi-iterations, can show very slow convergence.

In this work, we present a general family of solvers which leverage the fast converge of Krylov methods while being inherently inner-product free, and which are based off implicit LU factorizations of the matrix V_k in (2). These solvers construct bases that span the same Krylov subspaces as those associated to their traditional counterparts, but contain only linearly independent vectors that are not orthogonal by construction. The choice of the matrix-vector pair $\{C, d\}$ in (2) gives rise to different frameworks, which hold a strong parallelism to the Arnoldi and GKB algorithms: on the one hand, we use the standard Hessenberg method for $\{A, b\}$, see [4, 5], and on the other hand we develop a new modified version of the Hessenberg method for $\{A^T A, A^T b\}$ and for $\{AA^T, b\}$, see [7]. Note that, in practice, partial pivoting is needed to avoid unwanted break-down of the algorithms. Moreover, we describe the (quasi minimal residual) solvers associated with each approach. In the first place, we revisit the changing minimal residual Hessenberg method (CMRH), see e.g. [4, 5], and then we develop a completely new method, which we call the least squares LU (LSLU) [7]. In particular, we show that CMRH and LSLU can be used to tackle large-scale linear inverse problems efficiently, since they both present very fast convergence and regularization properties. Note that previous work on the CMRH method did not consider its application to ill-posed problems, and it was therefore not known if it was a regularizing method. Now we know that both CMRH and LSLU have empirical spectral filtering properties, i.e., early iterations reconstruct smooth component of the solution, and later iterations reconstruct high frequency components, so that early stopping of the iterations gives rise to a regularized solution. Finally, we extend both methods to include Tikhonov regularization, in the fashion of hybrid regularization [8], so that a projected Tikhonov regularization problem is solved at each iteration. This is a very powerful framework, since the solution of such problems does not show semi-convergence (and therefore is less sensitive to the stopping criteria), and we can choose the regularization parameters throughout the iterations using standard parameter choice criteria.

For more details on this work, where theoretical results and extensive numerical experiments suggest that inner-product free variants exhibit comparable performance to established methods, see [6, 7].

References

- [1] P. C. Hansen. Discrete Inverse Problems: Insight and Algorithms, SIAM (2010).
- [2] G. Meurant and J. Duintjer Tebbens, Krylov Methods for Nonsymmetric Linear Systems: From Theory to Computations, Springer Cham (2020).
- [3] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM (1994).
- [4] H. Sadok, CMRH: A new method for solving nonsymmetric linear systems based on the Hessenberg reduction algorithm, Numerical Algorithms, 20 (1999), pp.485-501.
- [5] H. Sadok and D. B. Szyld, A new look at CMRH and its relation to GMRES., BIT Numer. Math., 52 (2012), pp. 485–501.

- [6] A. N. Brown, J. G. Nagy and M. Sabaté Landman, H-CMRH: A Novel Inner Product Free Hybrid Krylov Method for Large-Scale Inverse Problems, Accepted for publication in SIMAX, (2024), arXiv:2401.06918.
- [7] A. N. Brown, J. Chung, J. G. Nagy and M. Sabaté Landman, *Inner Product Free Krylov Methods for Large-Scale Inverse Problems*, arXiv preprint, (2024), arXiv: arXiv:2409.05239.
- [8] J. Chung and S. Gazzola, Computational methods for large-scale inverse problems: a survey on hybrid projection methods, SIAM Review, 66 (2024), pp. 205–284.