

On efficiency and adaptivity of sketch-and-project approach in randomized linear solvers

Elizaveta Rebrova

Based on the joint works with Michal Dereziński, Deanna Needell, Daniel LeJeune, Jackie Lok.

Abstract

The sketch-and-project is a unifying framework for many known randomized iterative methods for solving linear systems, such as randomized Kaczmarz and coordinate descent algorithms, their block variants, as well as the extensions to non-linear optimization problems. Given a linear system $Ax = b$, the general scheme iterates to find its solution in the following way: for $t = 0, 1, \dots$ a random sketching matrix $S = S(t)$ is sampled from a distribution of (random) matrices and the update rule is given by

$$x_{t+1} = \arg \min_{x \in \mathbb{R}^n} \|x_t - x\|_B^2 \quad \text{such that} \quad SAx = Sb. \quad (1)$$

This optimization problem can be solved directly and is equivalent to an iterative step

$$x_{t+1} = x_t - B^{-1}A^\top S^\top (SAB^{-1}A^\top S^\top)^\dagger S(Ax_t - b).$$

The performance of these methods is often measured via the expected convergence rate:

$$E \|x_t - x_*\|_B^2 \leq (1 - \rho)^t \cdot \|x_0 - x_*\|_B^2 \quad \forall x_0, t$$

where $\rho \geq \lambda_{\min}(E[(S\tilde{A})^\dagger S\tilde{A}])$ and $\tilde{A} = AB^{-1/2}$. Some of well-known special cases, such as Randomized Kaczmarz method (when the sketching matrix S samples individual rows of A and $B = I$), have been mainly used as simple linear solvers that can demonstrate computational efficiency in the cases of vastly overdetermined linear systems, essentially, when the equations arrive in the streaming way. However, this perspective is far from complete. In the talk, I will discuss several recent results demonstrating other conceptual advantages of the sketch-and-project based linear solvers.

1. Fast linear solvers for the systems with low-rank structure.

It is expected that the sketching matrices $S \in \mathbb{R}^{k \times m}$ with larger sketch size k improve per-iteration convergence of the solver but clearly they become more computationally expensive within every step. In [DR24], we have quantified the advantage of increasing the sketch size and connected it to the *tail condition number* of the system, that is, singular values of A excluding the top k of them.

Based on this improved convergence rate analysis, one can build a practical linear solver with several very natural enhancements of the generic computation scheme, that is, (a) particular not too small sketch size and very sparse sketching, or block sampling on the system preconditioned by the Randomized Hadamard Transform (b) inexact solve in place of the pseudoinverse inversion, and (c) adding the momentum. Such solver can compute $\tilde{x}$ such that $\|A\tilde{x} - b\| \leq \epsilon \|b\|$ in time:

$\tilde{O}\left(\frac{\sigma_\ell}{\sigma_n} \cdot n^2 \log 1/\epsilon\right)$, where $\ell = C\sqrt{n}$ for some $C > 0$, where $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$ are singular values of the matrix A . This directly improves the standard time complexity for linear solvers, such

as $\tilde{O}(\kappa \cdot n^2)$, due to its independence from the full condition number (and the top of the spectrum in particular) on the linear systems with relatively flat tails of the spectrum. Such linear systems appear in a variety of standard applications such as spiked covariance models and kernel machines, or when the linear system is explicitly regularized, such as ridge regression.

Another consequence of these results, obtained in [DLNR24], is that the sketch-and-project approach can be also viewed as implicit preconditioning by the iterative sparse sketching (or sampling) procedure.

2. Subspace constrained iterative methods for linear solvers with prior information

The adaptive nature of the sketch-and-project iteration can be also used to guide the iterations when additional information is available about the system. In [LR24], we propose a version of the randomized Kaczmarz algorithm for solving systems of linear equations where the iterates are confined to the solution space of a selected subsystem. We show that the subspace constraint leads to an accelerated convergence rate, especially when the system has approximately low-rank structure that can be estimated before solving the system. Another natural place for a subspace constraint appears if only a part of a linear system changes with time or one is solving a sequence of otherwise connected linear systems. On Gaussian-like random data, we show that the proposed Subspace Constrained Randomized Kaczmarz method results in a form of dimension reduction that effectively increases the aspect ratio of the system.

3. Robust linear solvers for the systems with sparse corruptions

Finally, the adaptivity of the considered iterative framework gives another prominent application to solving linear systems contaminated with sparse corruptions. This setting models applications where some measurements are corrupted by arbitrarily large errors, which may occur during the data collection, transmission, and storage process due to malfunctioning sensors or faulty components. For sufficiently tall and regular systems $A \in R^{m \times n}$, $m \geq O(n)$, the iterates of a simple data-aware method based on the Randomized Kaczmarz algorithm converge to x_* avoiding a significant portion of arbitrary corruptions [HNRS22]. Moreover, the subspace constraining approach discussed above allows to efficiently utilize external knowledge about corruption-free equations and achieve convergence in difficult settings, such as not very overdetermined $((1 + \alpha)n \times n)$ linear systems [LR24].

References

- [DLNR24] Michał Dereziński, Daniel LeJeune, Deanna Needell, and Elizaveta Rebrova. Fine-grained analysis and faster algorithms for iteratively solving linear systems. *arXiv preprint arXiv:2405.05818*, 2024.
- [DR24] Michał Dereziński and Elizaveta Rebrova. Sharp analysis of sketch-and-project methods via a connection to randomized singular value decomposition. *SIAM Journal on Mathematics of Data Science*, 6(1):127–153, 2024.

- [HNRS22] Jamie Haddock, Deanna Needell, Elizaveta Rebrova, and William Swartworth. Quantile-based iterative methods for corrupted systems of linear equations. *SIAM Journal on Matrix Analysis and Applications*, 43(2):605–637, 2022.
- [LR24] Jackie Lok and Elizaveta Rebrova. A subspace constrained randomized kaczmarz method for structure or external knowledge exploitation. *Linear Algebra and its Applications*, 2024.