Riemannian optimization for matrix nearness problems

Vanni Noferini, Froilán Dopico, Miryam Gnazzo, Lauri Nyman, Federico Poloni

Abstract

Matrix nearness problems are central to numerical linear algebra and matrix theory. As highlighted in [6], these problems have wide-ranging applications. The proposed talk, outlined below, is based on several recent papers of mine, written with different coauthors [3, 4, 7].

Overview of matrix nearness problems

Let us start with a general description, inspired in part by Nick Higham's PhD thesis [5].

A matrix nearness problem involves finding a matrix B that is closest to a given matrix A, such that B satisfies a specific property $\mathfrak{P}$ which A does not. Formally, if $\mathcal{Q}$ represents the set of matrices that possess property $\mathfrak{P}$, the goal is to solve the following optimization problem:

$$\min \|A - X\| \quad \text{subject to} \quad X \in \mathcal{Q}. \tag{1}$$

The distance to be minimized in (1) is typically the Euclidean one, i.e., the distance induced by the Frobenius norm $\|X\|_F^2 = \text{tr}(X^*X)$, though other options can also be considered. Matrix nearness problems can be generalized to matrix pencils or matrix polynomials. For example, given a matrix polynomial $A(z) = \sum_{i=0}^d A_i z^i$, the objective becomes finding a polynomial $B(z) = \sum_{i=0}^d B_i z^i$ that minimizes the squared distance $\sum_{i=0}^d \|A_i - B_i\|^2$ while ensuring that $B(z)$ has the desired property $\mathfrak{P}$, which $A(z)$ lacks.

Some matrix nearness problems have well-understood, efficient solutions. For example, by the Eckart–Young–Mirsky theorem, the nearest singular matrix to a full-rank matrix $A \in GL(n, \mathbb{C})$ can be found via singular value decomposition (SVD). The solution is $B = \sum_{i=0}^{n-1} \sigma_i u_i v_i^*$, with the distance given by the smallest singular value σ_n.

However, many matrix nearness problems are more difficult, and have been either shown or conjectured to be NP-hard. The feasible set $\mathcal{Q}$ is often non-convex, making it challenging to find anything beyond a local minimum. Moreover, optimization algorithms that attempt the task are frequently quite slow and inefficient, making it almost impossible in practice to find even a local minimizer when the input size grows beyond very small matrices.

A new approach

In this talk, I will propose a new method to solve matrix nearness problems. There are three key features of the new approach:

1. It can handle a broad class of matrix nearness problems, including many described in the literature;

2. In extensive numerical experiments, the new method consistently outperforms existing algorithms, especially in challenging cases;

3. The problem is reformulated as an optimization task over Riemannian manifolds, offering a novel approach to previously intractable problems.

Our method is based on a key insight: many matrix nearness problems become more tractable when supplemented with additional information about the minimizer. This is akin to the concept of an "oracle" in theoretical computer science: an abstract machine that can solve specific problem instances in one step.

Let us denote the supplementary information about the optimizer by θ. The problem can then be stiffened by restricting the feasible set to matrices that share this information. Specifically, if we decompose the set $\mathcal{Q}$ as $\mathcal{Q} = \bigcup_\theta \mathcal{Q}_\theta$, where $\mathcal{Q}_\theta$ represents matrices with property $\mathfrak{P}$ and the additional attribute θ, we can solve the restricted problem:

$$f(\theta) = \min \|A - X\| \quad \text{subject to} \quad X \in \mathcal{Q}_\theta.$$

The original problem can then by equivalently reformulated by optimizing over θ:

$$\min_\theta f(\theta).$$

This often reduces the problem to an optimization task over a Riemannian manifold. For example, if θ is an eigenvector, the optimization is over the unit sphere. If θ represents a set of d independent eigenvectors, the optimization is over the Grassmann manifold of d-dimensional subspaces.

Of course, the idea to use Riemannian optimization to solve nearness problems is not new. Many works have applied manifold optimization to specific matrix nearness problems, but typically in a much more direct manner than the approach that I will present in this talk. To give but a couple of the many examples, Vandereycken [8] addressed low-rank matrix completion via optimization on fixed-rank manifolds, and Borsdorf [1] used augmented Lagrangian techniques for a chemistry-related problem over the Stiefel manifold. Oracle-based strategies have also previously appeared, such as the method by Byers [2] for finding the distance to Hurwitz instability.

However, our recent work has further advanced these ideas, and its distinctive feature is to use at the same time both oracles and Riemannian optimization, as well as other classical optimization tools such as regularization. It is, to my knowledge, the first time that this strategy has been attempted in a systematic way to tackle a wide range of matrix nearness problems.

A two-level optimization framework

More in detail, our method introduces a two-level optimization framework:

- **Inner Problem: Minimize the distance over $\mathcal{Q}_\theta$.** The inner subproblem is chosen so that it has a closed-form solution, or at least so that its solution can be computed cheaply.

- **Outer Problem: Minimize $f(\theta)$ over the manifold of all possible θ.** Riemannian optimization is then used to efficiently solve the outer problem.

We refer to this strategy as the *Riemann-Oracle* approach.

In the talk, I plan to describe how and why the Riemann-Oracle method can be applied to several matrix nearness problems of practical importance, including applications in numerical linear algebra, control theory, and computer algebra. Our experiments show that this approach consistently outperforms many specialized algorithms in both speed and accuracy.

I will explore in the talk both the theoretical details and the practical implementation aspects of this method. Furthermore, I will provide numerical experiments demonstrating concrete success stories. Examples include:

1. Finding the nearest matrix whose eigenvalues are all in a given set Ω (nearest stable matrix) [7]. Here the information θ is a unitary matrix that brings the solution into Schur form;

2. Finding the nearest singular pencil to a given one [3], including variants of the problem such as the nearest pencil with a prescribed minimal index. Here the information θ is a pair of unitary matrices that bring the solution into generalized Schur form;

3. Finding the nearest singular matrix with an additional linear constraint (e.g. sparsity pattern, Toeplitz structure, etc.) [4]. Here the information θ is an eigenvector;

4. Finding the nearest matrix of prescribed rank r and an additional linear constraint [4]. Here the information θ is a set of $n - r$ eigenvectors;

5. Finding the nearest unstable matrix, i.e., requiring at least one eigenvalue to lie outside the set Ω [4]. Here the information θ is an eigenvector;

6. Finding the nearest matrix polynomial of prescribed rank [4]. Here the information θ is a null polynomial vector;

7. Finding the approximate GCD of two scalar polynomials [4]. Here the information θ is a vector contatining the coefficients of two polynomials related to the problem.

References

[1] R. Borsdorf. An algorithm for finding an optimal projection of a symmetric matrix onto a diagonal matrix. *SIAM J. Matrix Anal. Appl.*, 35(1):198–224, 2014.

[2] R. Byers. A bisection method for measuring the distance of a stable matrix to the unstable matrices. *SIAM J. Sci. Stat. Comput.*, 9(5):875–881, 1988.

[3] F. Dopico, V. Noferini, and L. Nyman. A Riemannian optimization method to compute the nearest singular pencil. *SIAM J. Matrix Anal. Appl.*, To appear, 2024.

[4] M. Gnazzo, V. Noferini, L. Nyman, and F. Poloni. Riemann-Oracle: A general-purpose Riemannian optimizer to solve nearness problems in matrix theory. Preprint, 2024.

[5] N. J. Higham. *Nearness Problems in Numerical Linear Algebra.* PhD thesis, University of Manchester, 1985.

[6] N. J. Higham. Matrix nearness problems and applications. In *M. J. C. Gover and S. Barnett, editors, Applications of Matrix Theory*, pages 1–27. Oxford University Press, 1989.

[7] V. Noferini and F. Poloni. Nearest Ω-stable matrix via Riemannian optimization. *Numer. Math.*, 148:817–851, 2021.

[8] B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013.