Riemann-Oracle: A general-purpose Riemannian optimizer to solve nearness problems in matrix theory

Miryam Gnazzo, Vanni Noferini, Lauri Nyman, Federico Poloni

Abstract

Nearness problems are a fundamental class of problems in numerical linear algebra and matrix theory. A general description of matrix nearness problems is the following: given an input matrix A and a property $\mathfrak{P}$ that does not hold for A, the corresponding matrix nearness problem consists of finding a matrix B nearest to A and such that $\mathfrak{P}$ holds for B, as well as the distance between A and B. More formally, denoting by $\mathcal{Q}$ the set of matrices that have the property $\mathfrak{P}$, we seek the minimum and an argument minimum for the constrained optimization problem

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

In this work, we focus on nearness problems with respect to the Frobenius distance.

We propose an extremely versatile approach to address a large family of matrix nearness problems, possibly with additional linear constraints. Our method is based on splitting a matrix nearness problem into two nested optimization problems, of which the inner one can be solved either exactly or cheaply, while the outer one can be recast as an unconstrained optimization task over a smooth real Riemannian manifold. In the following, we illustrate this idea.

The method capitalizes on the insight that many matrix nearness problems become more tractable when provided by an oracle with specific supplementary information about the minimizer. For instance, if we are working over square matrices, this information could be a certain optimal eigenvalue or eigenvector, say, θ. Suppose further that, by restricting to matrices that share this information θ, the problem (1) can be solved exactly. In other words, let us write the feasible set as $\mathcal{Q} = \bigcup_\theta \mathcal{Q}_\theta$, where $\mathcal{Q}_\theta$ is the set of matrices having both the property $\mathfrak{P}$ and the attribute θ. Assume that we can compute the solution $f(\theta)$ of the stiffened problem

$$f(\theta) = \min \|A - X\| \quad \text{s.t. } X \in \mathcal{Q}_\theta. \tag{2}$$

At this point, we can solve the original problem (1) by optimizing (2) over θ, i.e., we can recast (1) as the equivalent optimization task

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

This minimization often has to be performed over an appropriate matrix or vector manifold; going back to the concrete example where θ is an eigenvector, and there are no additional restrictions on it, then one can optimize over the set of all possible normalized eigenvectors, i.e., the unit sphere. Or possibly, the information θ could consist of an arbitrary set of d linearly independent eigenvectors; in this case, one would then optimize over the Grassmannian of d-dimensional subspaces.

Nevertheless, we show that the objective function to be minimized on the Riemannian manifold can be discontinuous, thus requiring regularization techniques, and we give conditions for this to happen.

The ideas proposed in this work are applicable to a wide class of nearness problems. In essence, they amount to a two-level optimization framework, in which the inner subproblem can be solved cheaply and/or explicitly, while Riemannian optimization becomes a crucial tool for the outer subproblem.

We observe that this paradigm applies to many matrix nearness problems of practical interest appearing in the literature, thus revealing that they are equivalent in this sense to a Riemannian optimization problem. Among them, we recall the possible extension to nearness problem for matrix polynomials, as follows. In this case $A(z) = \sum_{i=0}^{d} A_i z^i$ is a polynomial with matrix coefficients, and one looks for a second matrix polynomial $B(z) = \sum_{i=0}^{d} B_i z^i$ that (a) minimizes the squared distance $\sum_{i=0}^{d} \|A_i - B_i\|^2$, and (b) has a given property $\mathfrak{P}$ that $A(z)$ lacks.

To illustrate the practical applicability of our approach, we consider the problems of finding the nearest singular polynomial matrix (given an upper bound on its degree), the approximate GCD (given a lower bound on its degree) of two scalar polynomials, and the nearest matrix with at least one eigenvalue in a prescribed closed set (distance to instability). We also include extensive numerical experiments, which demonstrate that our method often outperforms its predecessors, including algorithms specifically designed for those particular problems.

References

[1] Miryam Gnazzo, Vanni Noferini, Lauri Nyman and Federico Poloni (2024). *Riemann-Oracle: A general-purpose Riemannian optimizer to solve nearness problems in matrix theory*, `https://arxiv.org/abs/2407.03957`, arXiv.