3 January 2025 ·

Qeios

Peer Review

Review of: "Distributional Matrix Completion via Nearest Neighbors in the Wasserstein Space"

Matthew Werenski[1]

1. Computer Science, Tufts University, United States

The authors propose and study the problem of distributional matrix completion, which extends the well-studied problem of matrix completion with *scalar entries* to the setting where the matrix entries are *distributions over* $\mathbb{R}$. The approach considered uses a nearest-neighbor approach that can be broken down into two steps. For the ij'th entry

1. Find the rows that contain the column j and are closest to row i according to the average Wasserstein distance between the distributions in the shared columns.
2. Compute the Wasserstein barycenter of the distributions in column j using the rows that were identified as the closest in step 1.

This framework mirrors the standard nearest-neighbor approach to matrix completion but changes the measure of distance between pairs of rows as well as the averaging operation being used to estimate the missing entry. The identification of these changes opens the door for generalizing the framework to matrix completion over a manifold, as long as there are appropriate notions of distance and averaging.

In the setting they consider, one-dimensional distributions that satisfy certain structural conditions allow the authors to successfully derive asymptotic results for the convergence rates of the estimated distributions to their true distributions. The conditions imposed in Assumption 3 are common in the literature but are quite stringent, and a discussion of when they are satisfied would be welcomed. The results derived are made possible by considering the one-dimensional setting due to the closed form of the Wasserstein barycenter. Extending the results to higher-dimensional settings would likely

 doi.org/10.32388/2C2AQT

require a substantial amount of challenging analysis due to the complexity of Wasserstein barycenters in dimension 2 or more.

Extending the work into a more general setting would certainly be of interest, be it to this higher-dimensional setting or to a manifold framework. In either case, this work seems to open an interesting avenue of inquiry. It would be improved upon if the experimental section were furnished with a real-world application such as the teaching example discussed in the introduction, but the simulations are still interesting.

Declarations

Potential competing interests: No potential competing interests to declare.