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

Wenxing Zhang[1]

1 University of Electronic Science and Technology of China, China

Potential competing interests: No potential competing interests to declare.

Report on

Distributional matrix completion via nearest neighbors in the Wasserstein space

The authors proposed a very interesting generalization of matrix completion---distributional matrix completion. Instead of scalars as matrix entries, the 1D empirical distribution is considered as the matrix entries. The matching technique for large-scale matrix completion is transferred to the distributional matrix completion. The optimal transportation theory is exploited to measure the similarity of columns. The W2 distance is used. The Python code was rendered by the authors for reproducing the numerical experiments. Numerical results show that the proposed method not only recovers the unobserved distributions but also creates synthetic distributions that are consistently closer to the true distributions.

Only some comments and typos:

1. The authors discuss the “low-rank” of the distributional matrix. Although the authors provide some definition in the central paragraph of page 7, as a reader, I am still ambiguous about what the “low-rank” is. Can the authors provide a toy example for this new definition of the distributional matrix?
2. The authors use the matching technique for scalar matrix completions on the distributional matrix. Is it possible to develop some regularizer, e.g., the nuclear norm for scalar matrices, to regularize the low rankness of the distributional matrix? That would be more interesting than deploying the matching skill for distributional matrix completion.
3. What does “density is uniformly lower bounded and compact” mean?
4. Page 10, line 3, “We provide a proof of this theorem in”, the sentence seems to be not self-contained.
5. Page 10, line 9, “we can recover our the unobserved”, seems to have “our” removed.