Response to Reviews: JCGS-25-420

Jayani P. Gamage, Dianne Cook, Paul Harrison, Michael Lydeamore

2025-08-19

Reviewer 1

Summary:

This manuscript proposes an algorithm to evaluate nonlinear dimension reduction (NLDR) techniques. The authors identify some characteristics and limitations of NLDR and develop an algorithm to analyze NLDR results' quality of how well they match the original high-dimensional data.

Strengths:

1. This manuscript is well-organized, providing sufficient background knowledge and examples which makes this paper easy to follow.

2. This paper contains sufficient technical details in the introduction of the algorithm.

Weakness:

1. The main concern is whether the proposed algorithm can properly evaluate the NLDR's performance. Using the RMSE plot with different binwidth values (e.g., in Figure 8 and Figure 11) to evaluate NLDR's performance is problematic for not considering the density within each cluster. For NLDR results that have denser clusters, for example layout a,c,g,h in Figure 11, a smaller binwidth needs be used to achieve the same average count in each bin, which will cause their lines in Figure 11 to be more upper left. However, denser clusters are not worse NLDR results.

2. In Figure 7, as RMSE steadily increases with the binwidth, it's not clear why the three binwidths 0.03, 0.05, and 0.07 are chosen. In addition, we shouldn't use a single simulated dataset to determine the default hyperparameters without any real analysis.

3. For several figures that contains results from multiple NLDR algorithms, it would be nice to add the information of which algorithm/ what hyperparameter values for each of the subfigure. Many readers may have tried several NLDR algorithms themselves, and they will benefit from those figures to gain insights about behaviors of different NLDRs. Also it's nice to mention which data is used to generate each figure (e.g., in Figure 7).

4. In section 3.6.1, it says "Values of b1 between 2 and b1=sqrt(n/r2) are allowed", why the choice of b1 should consider r2?

Reviewer 2

Minor Issues

1. First part of the title seems inappropriate for a scientific paper to me, as does the implied (and not substantiated, IMO) claim that the proposal is (primarily) useful to reliably select one of many NLDRs (as opposed to primarily useful for visually investigating details about a specific 2D NLDR).

2. For marketing and clarity reasons, I would strongly suggest to not name the metric "RMSE", but something more distinctive/descriptive.

3. Flow from 3.1 to 3.2 (with too many sub-subsections) is difficult. Sections 3.2.1 (Scale the data) and 3.2.2 (Construct hexagon grid) are just short paragraphs, would be clearer if combined into a single numbered list of steps for the model construction. Much of 3 might be clearer and more compact if simply replaced by a direct step-by-step description of the algorithm in algorithmic pseudocode form with suitable short explanations.

4. Section 3.5 is a claim without supporting evidence – many out-of-sample embedding methods exist (some based on rather similar ideas, which should be referenced), you would need to show that this particular one actually works reasonably well for this to deserve space in the paper. I would strongly suggest to simply mention this in the discussion section as a possible avenue for further development instead.

5. Section 3.6 should be drastically shortened and details moved into an appendix.

6. p. 18: "A particular pattern that we commonly see is that analysts tend to pick layouts with clusters that have big separations between them." The literature/internet contains many lively discussions of the perils of mis-interpreting and post-hoc'ing NLDRs (e.g. (Izarry, 2024), (Chari & Pachter, 2023)), it might be good to cite/summarize these instead of/in addition to basing your motivation on personal anecdotal evidence? The tone here seems overly conversational to me (also: "we almost always see there are no big separations in the data") and should be made more formal.

7. Section 4, esp. p.19: IMO, "to compare and assess a range of representations" an analyst really should look at their respective RNX curves, stress values, and Shepard diagrams, think about whether they care more about local or global structure preservation, and then decide based on those factors. The method proposed here seems to me to be mostly useful to then investigate specific regions / slices of high-D space to see for which sets of data points the high-D structure is (not) preserved well in the NLDR. I find this whole section to be vastly overselling the proposal – it ignores and implicitly discards most of the prior work in this field. Its intro paragraphs are also highly redundant with other content in the paper and should be cut.

8. Section 5 should be moved into an appendix to shorten the paper somewhat. Figure in Section 5.1. could benefit from coloring the clusters in different colors so global structure preservation is also somewhat legible from the figures for the NLDRs (e.g. "is orange cluster between green and blue as in high-D or do they switch positions"). Section 5.2 might also have benefited from a corresponding ISOMAP or similar embedding of this data suitable for truthfully (isometrically!) unrolling the high-D manifold, and then showcasing that such a successful "unrolling" can be discerned from the way the projections look? tSNE just isn't isometric, so the distortions visible here are entirely expected, IMO?

Typos and grammar:

Revised version should be much more carefully proofread: "appropraite", "doen't", "it is has two separated nonlinear clusters", "the methods tNSE"