Human Guided Dimensionality Reduction

Christian Raue   Frederic Sadrieh

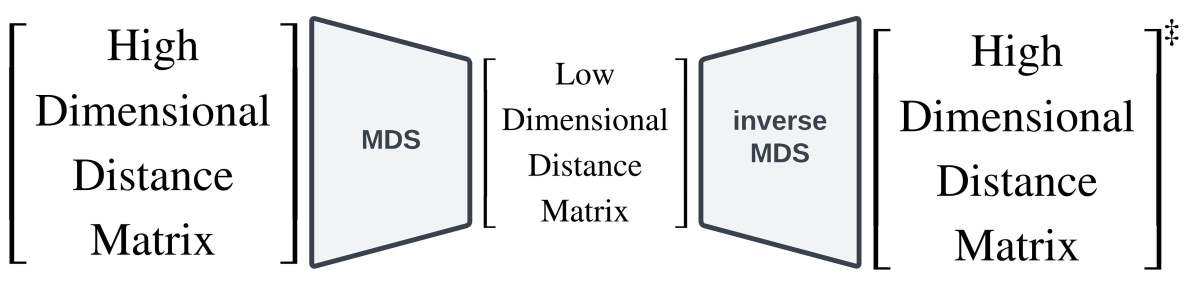

Figure 1: The Autoencoder nature of our inverse MDS.

Motivation

Dimensionality reduction (DR) allows humans to grasp high dimensional relationships between data points. DR is one-way, meaning that humans cannot influence how the high-dimensional data points are reduced [1]. We incorporate human interaction into DR. This allows the user to gain a better understanding of the DR and use their domain knowledge.

LMDS

Landmark Multi-Dimensional Scaling (LMDS) [2] is divided into two phases. Phase 1 processes a subset of points (the landmarks) using Multi-dimensional Scaling (MDS). MDS aims to preserve the similarity between the high and low dimensional distance matrices. The phase yields the Eigenvectors and Eigenvalues (EVVs) of the landmark distance matrix, in addition to the low dimensional landmark points. In phase 2 the remaining points are placed in low dimensional space based on the EVVs.

Our pipeline

We exploit this technique by incorporating human interaction (Fig. 2). First, the landmarks are randomly selected (1) and processed by MDS (2). Then the user can modify the low dimensional positions of the landmarks as desired (3). Inverse MDS transforms the new landmark positions back into a high dimensional distance matrix (4). Based on this matrix, new EVVs are generated (5). LMDS phase 2, projects the remaining points into low-dimensional space ((6) and (7)) and combines them with the low-dimensional landmarks (8). Metrics are then computed ((9) and (10)). They measure the quality of the dimensionality reduction. Finally, the human can reposition the landmarks based on these metrics and a scatter plot of the low-dimensional points (11). This creates a loop in which the human can understand the consequences of his actions and create an appropriate DR.

Steps (2) and (4) can be understood conceptually as an autoencoder (Fig. 1) (similar to [1]). With inverse MDS we want to invert the MDS, to do this, we train small neural networks with the Euclidean and Cosine distance metrics.

Inverse MDS

We train on the *imdb* and *emotion* datasets. As a hold-out task we use the *mnli* task from the *glue* dataset. We train a single layer neural network and optimize the hyperparameters using a Bayesian sweep.

Quantative Evaluation

How well can we invert MDS? How well can this be translated between datasets? We report the loss on the test sets of all our datasets (Table 2). The models achieve low test loss on their trained task. The transfer from the base task causes only small loss penalties, while the transfer between the distance metrics is much worse. This is to be expected, since cosine and euclidean have different ranges of values.

What impact does the landmark movement have? Using the metrics from [3], we were able to show that moving the landmarks can improve the quality of the DR (Table 1). Normalized stress suffers from landmark movement, since it is a measure of how similar the high and low distance matrices are, which is exactly what MDS optimizes.

Limitations

A qualitative user study with experts has not yet been conducted. We want to evaluate how users can improve their DR using our method. Our findings are limited by the use of only one sentence transformer model to embed our datasets. In addition, we train our inverse MDS models only on real MDS data, assuming that the MDS covers the whole low-dimensional space.

	initial landmarks	moved landmarks
Trustworthiness	0.384	0.588
Continuity	0.529	0.557
Normalized Stress	0.669	0.77
7-neighborhood-hit	0.5	0.556

Table 1: Metric values for LMDS with intial und moved landmarks

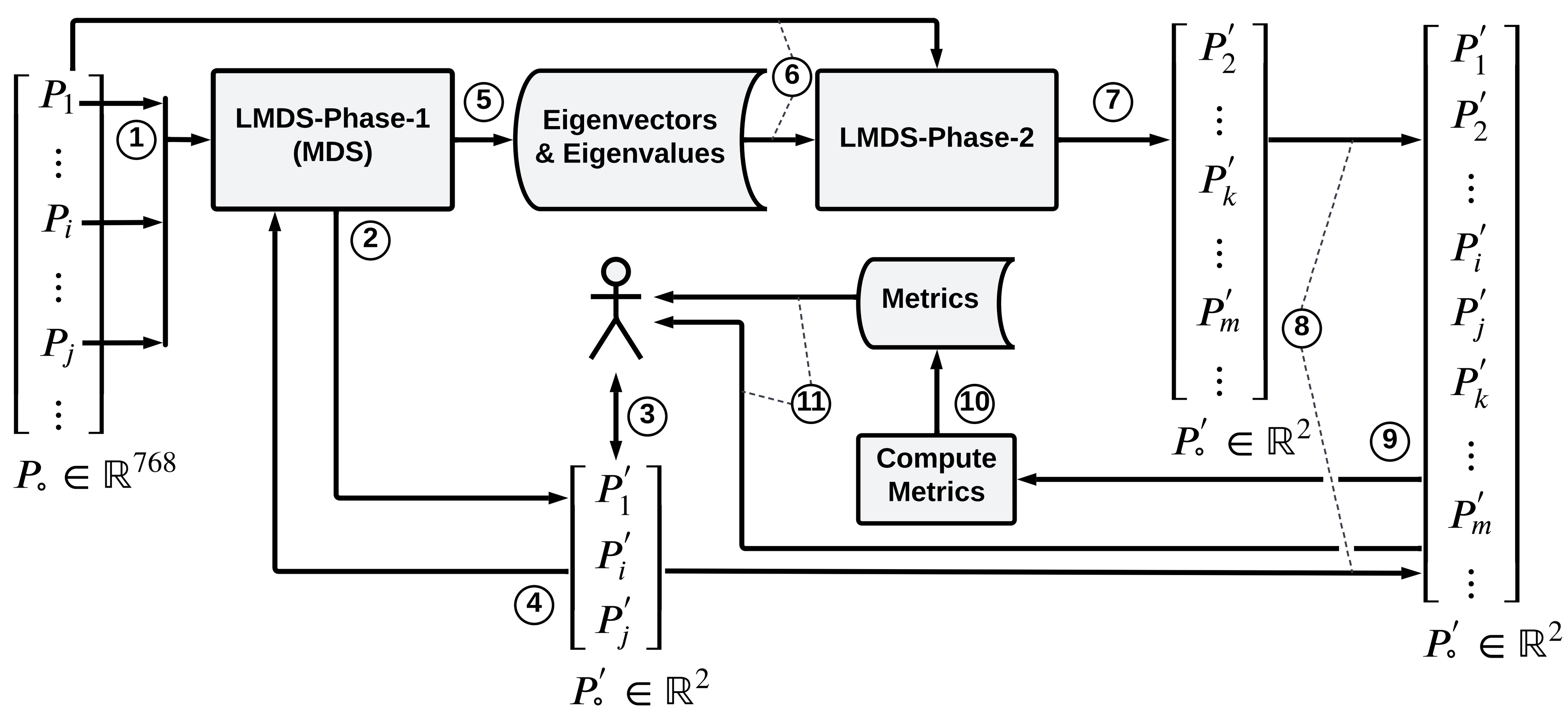

Figure 2: LMDS Process with human guidance.

	imdb_euclidean	imdb_cosine	emotion_euclidean	emotion_cosine	mnli_euclidean	mnli_cosine
imdb_euclidean	0.011	0.259	0.029	0.116	0.049	0.063
imdb_cosine	0.239	0.015	0.395	0.044	0.479	0.077
emotion_euclidean	0.033	0.407	0.008	0.216	0.009	0.139
emotion_cosine	0.105	0.046	0.21	0.013	0.271	0.016

Table 2: Loss of different models, based on their trained task (rows) performing inverse MDS on different unseen test datasets (columns)

References

M. Espadoto, G. Appleby, A. Suh, D. Cashman, M. Li, C. Scheidegger, E. W. Anderson, R. Chang, and A. C. Telea, "UnProjection: Leveraging inverse-projections for visual analytics of high-dimensional data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 2, pp. 1559–1572, 2023.

V. Silva and J. Tenenbaum, "Global Versus Local Methods in Nonlinear Dimensionality Reduction," in *Advances in Neural Information Processing Systems*, S. Becker, S. Thrun, and K. Obermayer, Eds., vol. 15. MIT Press, 2002. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2002/file/5d6646aad9bcc0be55b2c82f69750387-Paper.pdf

M. Espadoto, R. M. Martins, A. Kerren, N. S. T. Hirata, and A. C. Telea, "Toward a Quantitative Survey of Dimension Reduction Techniques," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 3, pp. 2153–2173, 2021.