Solving Large Scale Non-metric Multidimensional Scaling Problems Using ADMM Optimization

Tiffany Tsou and Ivy Zhang
Professor Lieven Vandenberghe, Xin Jiang, Taylor Chung

UCLA Samueli School of Engineering
SUMMER UNDERGRADUATE RESEARCH PROGRAM

INTRODUCTION + OBJECTIVE

The goal of this research project is to use an algorithm based on the alternating direction method of multipliers (ADMM), to solve large-scale non-metric multidimensional scaling (NMDS) problems with randomly generated datasets and more interesting and real-world datasets such as the Swiss roll, S curve, and related images. Our hypothesis is that the calculated Gram matrix will preserve the ordering of the original distances between points in our initial dataset.

PRINCIPLES

NMDS: attempts to preserve the original distances between inputs in a dataset.
ADMM: a method for large-scale optimization which performs alternating optimizations over two vector variables x and y.
Convex optimization: A convex objective function is subject to inequality constraints that are summarized by a slack variable that we seek to minimize.

Figures: S curve and Swiss roll.

Source: https://www.semanticscholar.org/paper/Nonlinear-Manifold-Learning-6-.-454-Summary-Ihler/62bc7f7507f8f3e7c9c4ac62215d31b06e45da98/figure/0

MATERIALS

Python: Optimization package cvxpy
Matlab: Optimization package CVX
Datasets: Random datasets of size 50, Swiss roll dataset, S curve dataset, images

METHODS

1. Generate datasets with random data points, Swiss roll, S curve, images.
2. Code optimization problem along with ADMM solution into Python and then into Matlab.
3. Analyze results in Python and Matlab.
4. Make necessary adjustments to code and repeat.

$$
\begin{array}{ll}
\min\limits_{K,\xi_{ijkl}} & \sum\limits_{(i,j,k,l)\in\mathcal{S}} \xi_{ijkl} + \lambda\,\mathrm{Trace}(K) \\
\text{subject to} & k_{kk} - 2k_{kl} + k_{ll} - k_{ii} + 2k_{ij} - k_{jj} \geq 1 - \xi_{ijkl} \\
& \sum\limits_{ab} k_{ab} = 0, \quad K \succeq 0. \qquad\qquad (\text{GNMDS})
\end{array}
$$

This is the optimization problem with inequality constraints represented as linear equations of Gram matrix K which define a unique K which can solve the problem and have specifications that disallow translations, rotations, and scalings of K.
Source: Agarwal et al.

$$
\begin{array}{ll}
\text{minimize} & \sum\limits_{k=1}^{m} \max\{0, u_k\} + \lambda \mathbf{1}^T x + g(x) + h(y) \\
\text{subject to} & \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} -I \\ -A \end{bmatrix} y = \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}.
\end{array}
$$

This is the optimization problem in an ADMM-ready form, which consists of first optimization over x and u, then optimization over y, and lastly the dual update. Source: Boyd et al.

RESULTS

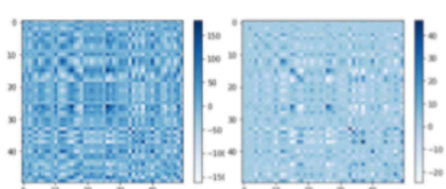
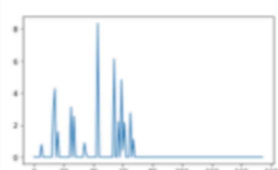


Figure: Shows the original Gram matrix calculated from the original distances matrix (left) and the calculated Gram matrix obtained by solving the non-metric multidimensional scaling problem (right).

Figure: Shows the values of the slack variable (error); ordering of distances is overall preserved.



REFERENCES

Agarwal, S., Wills, J., Cayton, L., Lanckriet, G., Kriegman, D., & Belongie, S. (2007, March). Generalized non-metric multidimensional scaling. In Artificial Intelligence and Statistics (pp. 11-18). PMLR.
Boyd, S., Parikh, N., & Chu, E. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. Now Publishers Inc.

ACKNOWLEDGEMENTS

We would like to thank the Summer Undergraduate Research Program for providing the guidance and resources to succeed with this project, Professor Lieven Vandenberghe, Xin Jiang, and Taylor Chung for providing their knowledge and resources, and NSF for funding this project.