Generalized Full-dimensional Scaling

Jan de Leeuw

First created on June 18, 2019. Last update on December 15, 2021

Abstract

If the n × p matrix X is a stationary point of the MDS loss function, then it is also the global minimum over the subspace of all n × p matrices with the same column space as X.

Contents

Note: This is a working paper which will be expanded/updated frequently. All suggestions for improvement are welcome. The directory deleeuwpdx.net/pubfolders/localglobal has a pdf version, the bib files, and the complete Rmd file.

1 Introduction

In (Euclidean, least squares, metric) multidimensional scaling (MDS) we minimize the *stress* loss function $\sigma(\bullet)$, defined as

$$\sigma(X)=\frac{1}{2}\sum_{1\leq i<j\leq n}w_{ij}(\delta_{ij}-d_{ij}(X))^2 \tag{1}$$

over all *configurations* $X\in\mathbb{R}^{n\times p}$, the linear space of $n\times p$ matrices.

Here $D(X)=\{d_{ij}(X)\}$ is a matrix of Euclidean distances between the rows of X, i.e.

$$d_{ij}(X)=\sqrt{\sum_{s=1}^p(x_{is}-x_{js})^2}.$$

We now introduce some standard MDS notation, following De Leeuw (1977). Define the unit vectors e_i, which have element i equal to one and all other elements equal to zero. For $i < j$ define the matrices

$$A_{ij} = (e_i - e_j)(e_i - e_j)'.$$

Note that $d_{ij}(X) = \sqrt{\text{tr } X'A_{ij}X}$. Next, define the matrix $V = \{v_{ij}\}$ by

$$V = \sum\sum_{1\leq i<j\leq n} w_{ij}A_{ij}. \tag{2}$$

Also define the matrix valued function $B(X) = \{b_{ij}(X)\}$ by

$$B(X) = \sum\sum_{1\leq i<j\leq n} w_{ij}r_{ij}(X)A_{ij} \tag{3}$$

where

$$r_{ij}(X) = \begin{cases} \frac{\delta_{ij}}{d_{ij}(X)} & \text{if } d_{ij}(X) > 0, \\ 0 & \text{if } d_{ij}(X) = 0. \end{cases}$$

We also assume, without loss of generality, that dissimilarities are normalized as

$$\frac{1}{2}\sum\sum_{1\leq i<j\leq n} w_{ij}\delta_{ij}^2 = 1.$$

Using these definitions and conventions gives

$$\sigma(X) = 1 - \text{tr } X'B(X)X + \frac{1}{2}\text{tr } X'VX,$$

and if $d_{ij}(X) > 0$ for all $i < j$

$$\mathcal{D}\sigma(X) = (V - B(X))X.$$

A configuration is a *stationary* point if $(V - B(X))X = 0$. A stationary point is *regular* if $d_{ij}(X) > 0$ for all $i < j$. De Leeuw (1984) shows that local minima are regular stationary points.

2 Main Result

Stationary points can be local minimum points or saddle points. The only local maximum point of stress is at $X = 0$ (De Leeuw (1993)). Among the local minimum points there are one or more global minimum points. A sufficient condition for a local minimum to be global is that at the stationary point we have $V - B(X) \gtrsim 0$, or $V^+B(X) \lesssim I$ (De Leeuw (2016)). This is a very restrictive condition which we generally do not expect to to be true. There is, however, a much weaker relation between stationary points and global minima on a subspace.

Theorem 1: [Local-Global] If $X \in \mathbb{R}^{n\times p}$ is a regular stationary point of the MDS problem then

$$\min_{T\in\mathbb{R}^{p\times p}} \sigma(XT) = \sigma(X).$$

Proof: First, observe that

$$d_{ij}(XT) = \sqrt{\text{tr } X'A_{ij}XTT'}$$

which is the square root of a non-negative linear function of $S = TT'$, and is consequently concave in S. It follows that

$$\sigma(XT) = 1 - \text{tr } X'B(XT)XS + \frac{1}{2}\text{tr } X'VXS$$

is convex in S. From Rockafellar (1970), theorem 31.4, the minimum over $S \gtrsim 0$ is attained at a unique point where

1. $S \gtrsim 0$.
2. $X'(V - B(XT))X \gtrsim 0$.
3. tr $X'(V - B(XT)X)XS = 0$.

But if X is a stationary point of the MDS problem we have $(V - B(X))X = 0$. Thus the minimum over S is attained at $S = I$, and the minimum over T is attained at any rotation matrix T with $T'T = TT' = I$, which is what the theorem says. ■

The part of theorem 1 where it is shown that $\sigma(XT)$ has a unique (and thus global) minimum over T for fixed X is mentioned in Borg and Groenen (2005), p 283. I merely added the result that the unique minimizer T is necessarily a rotation matrix if X is a stationary point. Note that the stationary point X may be a saddle point, it does not have to be a local minimum point.

An important special case of the theorem is full-dimensional scaling (De Leeuw (1993), De Leeuw, Groenen, and Mair (2016)), in which $p = n$.

Corollary 1: [Full] If $X \in \mathbb{R}^{n\times n}$ is a stationary point of the MDS problem then it is the unique global minimum.

Proof: In this case

$$\min_{T\in\mathbb{R}^{n\times n}} \sigma(XT) = \min_{Z\in\mathbb{R}^{n\times p}} \sigma(Z).$$

By theorem 1 consequently at a stationary point X

$$\sigma(X) = \min_{Z\in\mathbb{R}^{n\times p}} \sigma(Z).$$

■

References

Borg, I., and P. J. F. Groenen. 2005. *Modern Multidimensional Scaling*. Second Edition. Springer.

De Leeuw, J. 1977. "Applications of Convex Analysis to Multidimensional Scaling." In *Recent Developments in Statistics*, edited by J. R. Barra, F. Brodeau, G. Romier, and B. Van Cutsem, 133–45. Amsterdam, The Netherlands: North Holland Publishing Company.

———. 1984. "Differentiability of Kruskal's Stress at a Local Minimum." *Psychometrika* 49: 111–13.

———. 1993. "Fitting Distances by Least Squares." Preprint Series 130. Los Angeles, CA: UCLA Department of Statistics.

———. 2016. "Gower Rank." 2016.

De Leeuw, J., P. Groenen, and P. Mair. 2016. "Full-Dimensional Scaling." 2016.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press.