# Gower Rank

*Jan de Leeuw*

*Version 0.03, Jun 08, 2016*

**Abstract**

In Multidimensional Scaling we sometimes find that stress does not decrease if we increase dimensionality. This is explained in this note by using the Gower rank. Some examples with small Gower rank are analyzed.

## Contents

**Note:** This is a working paper which may be expanded/updated frequently. All suggestions for improvement are welcome. The directory gifi.stat.ucla.edu/gower has a pdf version, the complete Rmd file with all code chunks, and the bib file.

## 1 Theory

In *full-dimensional scaling* (De Leeuw (1993), De Leeuw, Groenen, and Mair (2016)) or *FDS* we minimize a loss function defined as

$$\sigma(C) := \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(\delta_{ij} - d_{ij}(C))^2 \tag{1}$$

over all positive semi-definite (psd) matrices $C$. Here $w_{ij}$ are known non-negative *weights* and $\delta_{ij}$ are known non-negative *dissimilarities*. To prevent complete triviality we assume $w_{ij}\delta_{ij} > 0$ for at least one pair $i, j$. The *distances* $d_{ij}(C)$ are defined by

$$d_{ij}(C) := \sqrt{c_{ii} + c_{jj} - 2c_{ij}}. \tag{2}$$

It follows that $\sigma$ is convex on the convex cone of psd matrices, and thus all local minimizers of the FDS problem are global. Suppose $\mathcal{I}$ is the set of all pairs $(i, j)$ with $w_{ij}\delta_{ij} > 0$. Then (De Leeuw (1984)) at a local minimum $d_{ij}(C) > 0$ for all $(i, j) \in \mathcal{I}$. Thus $\sigma$ is differentiable in a neighborhood of local minima.

We need some additional definitions. We use the *Loewner order*, i.e. $A \gtrsim B$ means that $A - B$ is psd. Define the matrices

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

with $e_i$ and $e_j$ unit vectors, i.e. vectors with one element equal to one and the rest equal to zero. Note that $d_{ij}(C) = \sqrt{\mathbf{tr}\ A_{ij}C}$. Using the matrices $A_{ij}$ we define

$$V := \sum_{j=1}^{n}\sum_{j=1}^{n} w_{ij} A_{ij}, \tag{3}$$

$$B(C) := \sum_{j=1}^{n}\sum_{j=1}^{n} w_{ij} \frac{\delta_{ij}}{d_{ij}(C)} A_{ij}. \tag{4}$$

If we assume, for convenience, that $\sum_{j=1}^{n}\sum_{j=1}^{n} w_{ij}\delta_{ij}^2 = 1$, then $\sigma(C) = 1 - 2\mathbf{tr}\ B(C)C + \mathbf{tr}\ VC$.

From convex analysis (Rockafellar (1970), theorem 31.4) the necessary and sufficient conditions for $C$ to be a minimizer are

$$C \gtrsim 0, \tag{5}$$
$$B(C) \lesssim V, \tag{6}$$
$$\mathbf{tr}\ (B(C) - V)C = 0. \tag{7}$$

In general we do not have uniqueness. Think of the case in which $\mathcal{I}$ has only one element, say $(1, 2)$. Then any psd matrix $C$ with $c_{11} + c_{22} - 2c_{12} = \delta_{12}^2$ is a minimizer with stress equal to zero. The local minimum is unique if and only if $C = 0$ is the only psd solution of the homogeneous linear system $d_{ij}^2(C) = \mathbf{tr}\ A_{ij}C = 0$ with $(i, j) \in \mathcal{I}$. Let us assume that this is the case, and that the global minimizer $C$ is of rank $p$. Then $p$ is called the *Gower rank* of the pair $(W, \Delta)$. If the minimum is not unique the Gower rank is the minimum rank of the minimizers.

There is a clear analogy with classical MDS, in which the corresponding quantity $p^\star$ is the number of positive eigenvalues of $-\frac{1}{2}J\Delta^{(2)}J$, the doubly-centered matrix of squared dissimilarities. In fact, the *Gower conjecture* is that the Gower rank $p$ is less than or equal to $p^\star$.

If $C = XX'$ is a full rank decomposition, with $X$ of dimensions $n \times p$, then we can write (7) as $V^+B(C)X = X$, using superscript $+$ for the Moore-Penrose inverse. Thus $X$ are eigenvectors of $V^+B(C)$ with eigenvalues equal to one. Condition (6) says that all eigenvalues of $V^+B(C)$ are less than or equal to one, and $X$ corresponds with the $p$ largest eigenvalues.

## 2   Algorithm

Multidimensional scaling (MDS) algorithms such as `smacof` (De Leeuw and Mair (2009)) parametrize the MDS problem in terms of an $n \times p$ *configuration* $X$. Thus

$$\sigma(X) := \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(\delta_{ij} - d_{ij}(X))^2 \tag{8}$$

with $d_{ij}(X) = \sqrt{\textbf{tr } A_{ij}XX'}$. An MDS(p) algorithm minimizes $\sigma$ over all $n \times p$ configurations.

An iterative MDS(p) algorithm converges to a configuration with $B(X)X = VX$, where $B(X)$ is defined by (4) with $C = XX'$. Thus $C = XX'$ solves (5) and (7), but the eigenvalues of $V^+B(X)$ can be larger than one, and consequently $C$ does not solve the FDS problem. But conversely if the MDS(p) solution $X$ gives a matrix $V^+B(X)$ with $p$ eigenvalues equal to one, then $p$ is the Gower rank and $C = XX'$ solves the FDS problem. This also implies that $X$ gives the global minimum of the MDS(q) problem for all $q \geq p$.

## 3   First Example

The following asymmetric matrix is taken from Wang and Chin (2010), table 5. Wong and Beasley (1990) constructed artificial data comparing seven hypothetical departments in a university in terms of three input variables (number of academic staff, academic staff salaries, support staff salaries) and three output variables (number of undergraduates, number of postgraduate students, number of resaerch papers). The data are used to construct a cross-efficiency matrix, which uses linear programming to compute input and output weights that measure efficiency of a department relative to another department. This produces cross efficiencies between zero and one, which can be interpreted as similarities, so we actually use one minus the data from Wang and Chin (2010).

```
##   a      b      b      d      d      f      g
## A 0.0000 0.0188 0.2310 0.3589 0.0618 0.0000 0.0000
## B 0.0781 0.0000 0.2281 0.2987 0.1010 0.0000 0.0000
## C 0.0000 0.1490 0.0000 0.5458 0.5050 0.0000 0.7059
## D 0.3125 0.0000 0.2651 0.1803 0.2350 0.0494 0.0000
## E 0.0000 0.1539 0.3349 0.5865 0.0000 0.0896 0.0000
## F 0.0000 0.0188 0.2310 0.3589 0.0618 0.0000 0.0000
## G 0.0000 0.0188 0.2310 0.3589 0.0618 0.0000 0.0000
```

It was observed by Shahin Ashkiani, PhD candidate at Autonomous University of Barcelona, that metric unfolding solutions for this matrix exhibit some curious behaviour (personal communication, June 4, 2016). Increasing the dimensionality of the solution does not decrease the minimum stress. The MDS(p) solution has the same minimum stress for all $p \geq 2$. This indicates the Gower rank of these data is two.

Although metric unfolding can be more efficiently done by `smacofRect()`, we use `smacofSym()` on the $14 \times 14$ augmented matrix, with the data in the off-diagonal submatrices and with zero weights for the diagonal submatrices. This corresponds more directly with the notation we use in this paper. We start with a random configuration and iterate until difference between successive stresses is less that 1e-15.

```
##  ndim = 1 | iterations =     4 | stress =   0.4693852450

##  ndim = 2 | iterations =   150 | stress =   0.4067723531

##  ndim = 3 | iterations =   133 | stress =   0.4067723531

##  ndim = 4 | iterations =   142 | stress =   0.4067723531

##  ndim = 5 | iterations =   113 | stress =   0.4067723531

##  ndim = 6 | iterations =   135 | stress =   0.4067723531
```

We can also illustrate this by showing the singular values of the configuration computed by `smacofSym()`.

```
## [1] 3.099044

## [1] 2.574971 1.997667

## [1] 2.5749717 1.9976662 0.0000001

## [1] 2.574971 1.997667 0.000000 0.000000

## [1] 2.5749716 1.9976663 0.0000003 0.0000000 0.0000000

## [1] 2.574971 1.997667 0.000000 0.000000 0.000000 0.000000
```

# 4  Second Example

There is a more interesting example in Wang and Chin (2010), analyzing cross efficiencies computed from data of Tofallis (1997). In this case the decision making units are 14 airlines. The input variables are aircraft capacity, operating cost, and non-flight assets. The output variables are passenger kilometres and non-passenger revenue. The cross-efficiencies are

```
##   a    b    c    d    e    f    g    h    i    j    k    l
## A 0.13 0.55 0.38 0.13 0.15 0.53 0.19 0.21 0.30 0.25 0.13 0.23
## B 0.83 0.66 0.95 0.83 0.83 0.98 0.75 0.73 0.72 0.79 0.83 0.80
## C 0.12 0.81 0.05 0.12 0.12 0.31 0.28 0.32 0.38 0.22 0.12 0.19
```

```
## D   0.04   0.57   0.30   0.04   0.06   0.30   0.18   0.21   0.30   0.19   0.04   0.17
## E   0.03   0.63   0.00   0.03   0.00   0.00   0.23   0.26   0.22   0.00   0.03   0.00
## F   0.12   0.89   0.04   0.12   0.12   0.02   0.34   0.39   0.49   0.28   0.12   0.25
## G   0.08   0.22   0.52   0.08   0.12   0.66   0.00   0.00   0.16   0.22   0.08   0.20
## H   0.22   0.39   0.48   0.22   0.23   0.71   0.15   0.14   0.18   0.25   0.22   0.24
## I   0.21   0.27   0.49   0.21   0.21   0.73   0.12   0.09   0.05   0.16   0.21   0.16
## J   0.22   0.36   0.35   0.22   0.17   0.64   0.22   0.21   0.00   0.00   0.22   0.03
## K   0.00   0.00   0.57   0.00   0.00   0.56   0.00   0.00   0.00   0.00   0.00   0.00
## L   0.05   0.37   0.25   0.05   0.04   0.56   0.06   0.06   0.00   0.00   0.05   0.00
## M   0.00   0.57   0.00   0.00   0.00   0.54   0.00   0.00   0.00   0.02   0.00   0.00
## N   0.00   0.77   0.00   0.00   0.00   0.00   0.22   0.27   0.35   0.14   0.00   0.12
##      m      n
## A   0.13   0.13
## B   0.83   0.83
## C   0.12   0.12
## D   0.04   0.04
## E   0.03   0.03
## F   0.12   0.12
## G   0.08   0.08
## H   0.22   0.22
## I   0.21   0.21
## J   0.22   0.22
## K   0.00   0.00
## L   0.05   0.05
## M   0.00   0.00
## N   0.00   0.00
```

We repeat the same analysis as in the first example. In this case we conclude the Gower rank is equal to three.

```
##   ndim = 1 | iterations =      6  | stress =   0.4760530537

##   ndim = 2 | iterations =    278  | stress =   0.2437571250

##   ndim = 3 | iterations =   1523  | stress =   0.2386715692

##   ndim = 4 | iterations =   1943  | stress =   0.2386715692

##   ndim = 5 | iterations =   2013  | stress =   0.2386715692

##   ndim = 6 | iterations =   1862  | stress =   0.2386715692
```

The singular values for the second example are

```
## [1] 4.503339

## [1] 4.153864 2.830820

## [1] 4.151090 2.632152 1.042914

## [1] 4.151090 2.632154 1.042909 0.000020
```

```
## [1] 4.151090 2.632154 1.042909 0.000020 0.000000
## [1] 4.151090 2.632154 1.042909 0.000020 0.000000 0.000000
```

# 5 Third Example

The third example was already analyzed in De Leeuw, Groenen, and Mair (2016). The Ekman (1954) color data give similarities between 14 colors. We transform to dissimilarities by using the transformation $(1 - x)^3$, and we repeat the analysis from the previous sections. Note that this is not an unfolding example, all weights are equal to one. The Gower rank turns out to be two, i.e. for this transformation the two-dimensional `smacof` solution is the global minimizer of stress in $p \geq 2$ dimensions.

```
##  ndim = 1 | iterations =     7  | stress =  0.5131738031
##  ndim = 2 | iterations =    44  | stress =  0.1049991045
##  ndim = 3 | iterations =   161  | stress =  0.1049991045
##  ndim = 4 | iterations =   178  | stress =  0.1049991045
##  ndim = 5 | iterations =   167  | stress =  0.1049991045
##  ndim = 6 | iterations =   172  | stress =  0.1049991045
```

The singular values for the third example are

```
## [1] 2.188205
## [1] 2.060186 1.477827
## [1] 2.0601856 1.4778274 0.0000007
## [1] 2.0601856 1.4778274 0.0000007 0.0000000
## [1] 2.0601856 1.4778274 0.0000007 0.0000000 0.0000000
## [1] 2.0601856 1.4778274 0.0000007 0.0000000 0.0000000 0.0000000
```

# 6 NEWS

001 06/06/16

- First release

002 08/06/16

- Third Example

003 08/06/16

- Added something on uniqueness
- Added Gower conjecture

# References

De Leeuw, J. 1984. "Differentiability of Kruskal's Stress at a Local Minimum." *Psychometrika* 49: 111–13. http://www.stat.ucla.edu/~deleeuw/janspubs/1984/articles/deleeuw_A_84f.pdf.

———. 1993. "Fitting Distances by Least Squares." Preprint Series 130. Los Angeles, CA: UCLA Department of Statistics. http://www.stat.ucla.edu/~deleeuw/janspubs/1993/reports/deleeuw_R_93c.pdf.

De Leeuw, J., and P. Mair. 2009. "Multidimensional Scaling Using Majorization: SMACOF in R." *Journal of Statistical Software* 31 (3): 1–30. http://www.stat.ucla.edu/~deleeuw/janspubs/2009/articles/deleeuw_mair_A_09c.pdf.

De Leeuw, J., P. Groenen, and P. Mair. 2016. "Full-Dimensional Scaling." doi:10.13140/RG.2.1.1038.4407.

Ekman, G. 1954. "Dimensions of Color Vision." *Journal of Psychology* 38: 467–74.

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

Tofallis, C. 1997. "Input Efficiency Profiling: an Application to Airlines." *Computers and Operations Research* 24: 253–58.

Wang, Y.-M., and K.-S. Chin. 2010. "A Neutral DEA Model for Cross-efficiency Evaluation and its Extension." *Expert Systems with Applications* 37: 3666–75.

Wong, Y.-H. B., and J.E. Beasley. 1990. "Restricting Weight Flexibility in Data Envelopment Analysis." *The Journal of the Operational Research Society* 41: 829–35.