

Differentiability of rStress at a Local Minimum

Jan de Leeuw, Patrick Groenen, Patrick Mair

Version 003, February 8, 2016

Contents

1	Problem	1
2	Directional Derivatives	2
3	Results	2
4	Local Maximum	3
5	NEWS	3
	References	3

Note: This is a working paper which will be expanded/updated frequently. The directory gifi.stat.ucla.edu/rstressdiff has a pdf copy of this article and the complete Rmd file.

1 Problem

We study differentiability of the multidimensional scaling loss function rStress (De Leeuw (2016)), defined as

$$\sigma_r(x) := \sum_{i=1}^n w_i (\delta_i - (x' A_i x)^r)^2 \quad (1)$$

for some $r > 0$. Here the w_i are positive weights and the δ_i are positive dissimilarities. The matrices A_i are positive semi-definite, and the quantities $x' A_i x$ are squared distances.

Clearly if $x' A_i x > 0$ for all i the loss function is differentiable. De Leeuw (1984) proves directional differentiability for $r = \frac{1}{2}$ and he shows that at a local minimum we generally have $x' A_i x > 0$. We investigate if and how this results generalizes to σ_r .

2 Directional Derivatives

Define the directional derivative

$$d\sigma_r(x, y) := \lim_{\epsilon \downarrow 0} \frac{\sigma_r(x + \epsilon y) - \sigma_r(x)}{\epsilon}.$$

For our computations we need

$$\begin{aligned} I_+(x) &:= \{i \mid x' A_i x > 0\}, \\ I_0(x) &:= \{i \mid x' A_i x = 0\}. \end{aligned}$$

Then

$$\begin{aligned} \frac{\sigma_r(x + \epsilon y) - \sigma_r(x)}{\epsilon} &= -4r \sum_{i \in I_+} w_i (\delta_i - (x' A_i x)^r) (x' A_i x)^{r-1} y' A_i x \\ &\quad - 2\epsilon^{2r-1} \sum_{i \in I_0} w_i \delta_i (y' A_i y)^r + \epsilon^{4r-1} \sum_{i \in I_0} w_i (y' A_i y)^{2r} + \frac{o(\epsilon)}{\epsilon}, \end{aligned}$$

and thus

$$d\sigma_r(x, y) = \begin{cases} -4r \sum_{i=1}^n w_i (\delta_i - (x' A_i x)^r) (x' A_i x)^{r-1} y' A_i x & \text{if } r > \frac{1}{2}, \\ -4r \sum_{i \in I_+} w_i (\delta_i - (x' A_i x)^r) (x' A_i x)^{r-1} y' A_i x - 2 \sum_{i \in I_0} w_i \delta_i (y' A_i y)^r & \text{if } r = \frac{1}{2}, \\ +\infty & \text{if } r < \frac{1}{2}. \end{cases}$$

3 Results

From our computations we derive the following results.

Theorem 1: If $r > \frac{1}{2}$ then σ_r is differentiable at x . If σ_r has a local minimum at x then

$$\sum_{i=1}^n w_i \delta_i (x' A_i x)^{r-1} A_i x = \sum_{i=1}^n w_i (x' A_i x)^{2r-1} A_i x.$$

Theorem 2: If $r = \frac{1}{2}$ then σ_r is directionally differentiable at x in every direction y . If σ_r has a local minimum at x then

$$\sum_{i \in I_+(x)} w_i \delta_i (x' A_i x)^{r-1} A_i x = \sum_{i \in I_+(x)} w_i (x' A_i x)^{2r-1} A_i x.$$

and $I_0(x) = \emptyset$.

Theorem 3: If $r < \frac{1}{2}$ then σ_r is directionally differentiable only in those directions y with $y' A_i y = 0$ for all $i \in I_0(x)$.

Thus for $r = \frac{1}{2}$ we have non-zero distances and differentiability at local minima, for $r > \frac{1}{2}$ it is quite possible that local minima with zero distances exist, and for $r > \frac{1}{2}$ rStress is not even directionally differentiable at points with zero distances.

4 Local Maximum

We can also generalize a result of De Leeuw (1993) to rStress.

Theorem 4: σ_r has a local maximum at x if and only if $x = 0$.

Proof: If $x = 0$ then

$$\sigma_r(x + \epsilon y) - \sigma_r(x) = -2\epsilon^{2r} \left\{ \sum_{i=1}^n w_i \delta_i (y' A y)^r - \frac{1}{2} \epsilon^{2r} \sum_{i=1}^n w_i (y' A_i y)^{2r} \right\}.$$

It follows that if

$$\frac{1}{2} \epsilon^{2r} \leq \frac{\sum_{i=1}^n w_i \delta_i (y' A y)^r}{\sum_{i=1}^n w_i (y' A_i y)^{2r}}$$

we have $\sigma(x + \epsilon y) - \sigma(x) \leq 0$. So, although σ_r may not even directionally differentiable at $x = 0$, it does decrease in all directions and is thus a local minimum.

Converse, suppose σ_r has a local maximum at $x \neq 0$. Then

$$\sigma_r(\epsilon x) = \sum_{i=1}^n w_i \delta_i^2 - 2\theta \sum_{i=1}^n w_i \delta_i (x' A x)^r + \theta^2 \sum_{i=1}^n w_i (x' A_i x)^{2r},$$

with $\theta := \epsilon^{2r}$. Thus σ_r is a convex quadratic in θ and it cannot have a local maximum on the ray through x . **QED** \$\$

5 NEWS

001 01/14/16 – First upload

002 01/15/16 – Added local maximum result

003 02/08/16 – Corrected some typos

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.

———. 2016. “Minimizing rStress Using Majorization.” <http://rpubs.com/deleeuw/142619>.