Higher Partials of Stress. Who Needs Them ?

Jan de Leeuw

Version 01, July 20, 2017

Abstract

Third

Contents

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

1 Introduction

2 Partials, Partials Everywhere

2.1 Derivatives of Distances

We start with the approximation

$$\frac{\sqrt{x+\epsilon}}{\sqrt{x}}=1+\frac{1}{2}\left(\frac{\epsilon}{x}\right)-\frac{1}{8}\left(\frac{\epsilon}{x}\right)^2+\frac{1}{16}\left(\frac{\epsilon}{x}\right)^3-\frac{5}{128}\left(\frac{\epsilon}{x}\right)^4+o(\epsilon^4).$$

Now let $e(x):=x'Ax$ and define

$$h(x,y):=\frac{\sqrt{e(x+y)}}{\sqrt{e(x)}}=\frac{\sqrt{x'Ax+2x'Ay+y'Ay}}{\sqrt{x'Ax}}.$$

Then

$$h(x,y)=1+\frac{1}{2}\left(\frac{2x'Ay+y'Ay}{x'Ax}\right)-\frac{1}{8}\left(\frac{2x'Ay+y'Ay}{x'Ax}\right)^2+\frac{1}{16}\left(\frac{2x'Ay+y'Ay}{x'Ax}\right)^3 -\frac{5}{128}\left(\frac{2x'Ay+y'Ay}{x'Ax}\right)^4+o(\|y\|^4).$$

If we expand powers and collect terms we get, assuming we have been lucky,

$$h(x,y)=1+h_1(x,y)+\frac{1}{2}h_2(x,y)+\frac{1}{6}h_3(x,y)+\frac{1}{24}h_4(x,y)+o(\|y\|^4)$$

$$h_1(x,y)=\frac{x'Ay}{x'Ax}$$

$$h_2(x,y)=\frac{y'Ay}{x'Ax}-\frac{(x'Ay)^2}{(x'Ax)^2}$$

$$h_3(x,y)=3\left\{\frac{(x'Ay)^3}{(x'Ax)^3}-\frac{(x'Ay)(y'Ay)}{(x'Ax)^2}\right\}$$

$$h_4(x,y)=-3\frac{(y'Ay)^2}{(x'Ax)^2}+18\frac{(x'Ay)^2y'Ay}{(x'Ax)^3}-15\frac{(x'Ay)^4}{(x'Ax)^4}$$

2.2 Derivatives of Stress

The stress loss function is

$$\sigma(x):=\frac{1}{2}\mathop{\sum\sum}_{1\leq i<j\leq n}w_{ij}(\delta_{ij}-d_{ij}(x))^2,$$

Here $x=\mathbf{vec}(X)$, with X the usual MDS configuration of n points in p dimensions. The distances between points i and j in the configurations are $d_{ij}(x):=\sqrt{x'A_{ij}x}$.

The $np\times np$ matrices A_{ij} are defined using unit vectors e_i and e_j, all zero except for one element that is equal to one. If you like, the e_i are the columns of the identity matrix. Define

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

and use p copies of E_{ij} to make the direct sum

$$A_{ij}:=\underbrace{E_{ij}\oplus\cdots\oplus E_{ij}}_{p\text{ times}}.$$

Assuming without loss of generality that the dissimilarities are normalized to sum of squares one. Then

$$\sigma(x) = 1 - \rho(x) + \frac{1}{2}x'Vx,$$

with

$$\rho(x) := \mathop{\sum\sum}_{1\leq i<j\leq n} w_{ij}\delta_{ij}\sqrt{x'A_{ij}x},$$

and

$$V := \mathop{\sum\sum}_{1\leq i<j\leq n} w_{ij}A_{ij}.$$

The function $\frac{1}{2}x'Vx$ is quadratic, so its derivatives are trivial to compute.

$$\rho(x+y) = \mathop{\sum\sum}_{1\leq i<j\leq n} w_{ij}\delta_{ij}\sqrt{e_{ij}(x+y)}$$

$$\rho(x+y) = \rho(x) + y'\mathcal{D}\rho(x) + \frac{1}{2}y'\mathcal{D}^2(x)y +$$

$$\mathcal{D}\rho(x) = \mathop{\sum\sum}_{1\leq i<j\leq n} w_{ij}\frac{\delta_{ij}}{d_{ij}(x)}A_{ij}x$$

$$\mathcal{D}^2\rho(x) = \mathop{\sum\sum}_{1\leq i<j\leq n} w_{ij}\delta_{ij}$$