Jan de Leeuw

34

SMOOTHNESS PROPERTIES OF NONMETRIC LOSS FUNCTIONS

Jan de Leeuw*
Bell Laboratories
Murray Hill, New Jersey 07974

ABSTRACT

The smoothness of some of the more important loss functions used in nonmetric scaling is investigated. We prove the following results

Loss Function	Continuous	Differentiable	Twice Differentiable
Kruskal's Stress	Yes	Yes	No
Guttman's Phi	Yes	No	No

*On leave of absence from the University of Leiden, Leiden, The Netherlands.

SMOOTHNESS PROPERTIES OF NONMETRIC LOSS FUNCTIONS

Jan de Leeuw*
Bell Laboratories
Murray Hill, New Jersey 07974

0: Introduction

In this paper we give some simple examples concerning behavior of Kruskal's stress and Guttman's phi in the case of a linear measurement model. The examples can be used for teaching purposes, but they also illustrate some general facts about continuity and differentiability of the loss functions.

1: A First Example

Consider the linear model

$$\begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} -\theta_1-\theta_2 \\ \theta_1-\theta_2 \\ -\theta_1+\theta_2 \\ \theta_1+\theta_2 \end{bmatrix}$$

We desire

$$Y_1 \le Y_2 \le Y_3 \le Y_4.$$

Without loss of generality we restrict ourselves to those θ satisfying

---

*On leave of absence from the University of Leiden, Leiden, The Netherlands.

$$\theta_1^2 + \theta_2^2 = 1$$

We define

$$S^*(\theta) = \frac{1}{16}\{(Y_1-Y_1^*)^2 + (Y_2-Y_2^*)^2 + (Y_3-Y_3^*)^2 + (Y_4-Y_4^*)^2\},$$

with Y* a permutation of Y satisfying

$$Y_1^* \leq Y_2^* \leq Y_3^* \leq Y_4^*.$$

We also define

$$S^0(\theta) = \frac{1}{4}\{(Y_1-Y_1^0)^2 + (Y_2-Y_2^0)^2 + (Y_3-Y_3^0)^2 + (Y_4-Y_4^0)^2\},$$

with Y^0 the (unique) vector satisfying

$$Y_1^0 \leq Y_2^0 \leq Y_3^0 \leq Y_4^0$$

and minimizing $S^0(\theta)$ for fixed Y. Both coefficients vary in the closed interval [0,1].

Moreover,

$$\begin{aligned} S^0(\theta) = 0 \quad \text{iff } S^*(\theta) = 0 \quad & \text{iff } Y^0 = Y^* = Y. \\ & \text{iff } Y_1 \leq Y_2 \leq Y_3 \leq Y_4. \\ & \text{iff } 0 \leq \theta_1 \leq \theta_2 \end{aligned}$$

and

$$S^*(\theta) = 1 \ \text{ iff } Y^* = -Y \ \text{ iff } Y_4 \le Y_3 \le Y_2 \le Y_1$$
$$\text{iff } \theta_2 \le \theta_1 \le 0$$

and

$$S^0(\theta) = 1 \ \text{ iff } Y^0 = 0 \ \text{ iff } Y_{[2,3,4]} \le Y_{[1]}$$
$$\wedge \ Y_{[3,4]} \le Y_{[1,2]} \wedge Y_{[4]} \le Y_{[1,2,3]}$$
$$\text{iff } \theta_2 \le 0 \wedge \theta_1 + \theta_2 \le 0$$

We use the notation $Y_{[i_n,\ldots,i_p]}$ for the average of $Y_{i_1},\ldots,Y_{i_p}$. Table 1:1 gives the different regions in (θ_1,θ_2) space with their different characterizations. Figure 1:2 gives a plot of $S^*(\theta)$ using polar coordinates, and 1:3 gives a similar plot of $S^0(\theta)$. From the tables and figures it is clear that both functions are continuous. Moreover, $S^0(\theta)$ is continuously differentiable, while the derivatives of $S^*(\theta)$ do not exist at the points $\frac{1}{2}\pi$, $\frac{3}{4}\pi$, $\frac{3}{2}\pi$, and $\frac{7}{4}\pi$. The second derivatives of $S^*(\theta)$ do not exist at all eight "special" points and those of $S^0(\theta)$ do not exist at $\frac{1}{4}\pi$, $\frac{3}{4}\pi$, $\frac{3}{2}\pi$, and 2π. The concave appearance of the function is deceptive; if we had broken open the circle at π, then the one-dimensional plot would be convex-like. Of course, this example is unrealistic for various reasons.

In the first place, a perfect solution exists. In the second place, there are no real local minima except the platform of $S^*(\theta)$ between $\frac{1}{2}\pi$ and $\frac{3}{4}\pi$ and between $\frac{3}{4}\pi$ and $\frac{7}{4}\pi$. In real examples, perfectly flat areas like this usually will not occur. There may, of course, be a considerable region with $S^*(\theta) = 1$ and/or $S^0(\theta) = 1$.

2: <u>A Second Example</u>

It is clear from Table 1:1 that the linear model

$$\begin{bmatrix} -1 & -1 \\ 1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} -\theta_1-\theta_2 \\ \theta_1+\theta_2 \\ -\theta_1+\theta_2 \\ \theta_1-\theta_2 \end{bmatrix}$$

with requirements

$$Y_1 \le Y_2 \le Y_3 \le Y_4$$

does not have a perfect solution. The relevant information is contained in Table 2:1, and the plots are drawn in 2:2 and 2:3.

The results are quite different in several respects. In the first place, there are a number of stationary values. The functions $S^*(\theta)$ has local maxima at $\frac{1}{2}\pi$, $\frac{5}{4}\pi$, $\frac{7}{4}\pi$, local minima at $\frac{1}{4}\pi$, $\frac{3}{4}\pi$, $\frac{3}{2}\pi$, and saddle values at π, 2π. The function is not differentiable at all these minima and maxima. The function $S^0(\theta)$ is more smooth. It has a local maximum at $\frac{1}{2}\pi$, and two local minima at $\frac{1}{4}\pi$ and $\frac{3}{4}\pi$ (both local minima are also global). It is continuously differentiable at all points, but at the special points the second order partials do not exist.

3: Some General Considerations

The examples in the previous sections illustrate some general theorems. Define the following general versions of the loss functions (compare Kruskal, 1964 a,b, 1965; Guttman, 1968).

$$S^0(\theta) = \frac{\sum_{i=1}^{n} (\phi_i(\theta)-\phi_i^0)^2}{\sum_{i=1}^{n} \phi_i^2(\theta)}$$

$$S^*(\theta) = \frac{\sum_{i=1}^{n} (\phi_i(\theta)-\phi_i^*)^2}{\sum_{i=1}^{n} \phi_i^2(\theta)}$$

Here $\phi_i(\theta)$ is defined by the measurement model we are considering. We assume that $\phi_i(\theta)$ is a continuously differentiable function of $\theta\varepsilon\Theta$. As usual ϕ_i^0 is defined as a set of numbers, satisfying the order restrictions in the data, and minimizing S^0 for fixed $\phi_i(\theta)$, and ϕ_i^* is defined as a permutation of the $\phi_i(\theta)$, satisfying the order restrictions, and minimizing S^* for fixed $\phi_i(\theta)$. In general, ϕ_i^0 solves a quadratic programming problem of a special structure, and finding ϕ_i^* is equivalent to solving a linear programming problem with a special structure. It easily follows from this remark that both $S^0(\theta)$ and $S^*(\theta)$ are continuous functions of θ for all points for which the denominator does not vanish. In fact, we can prove more. It also follows from this representation that the numerators of $S^0(\theta)$ and $S^*(\theta)$ are convex functions of $\phi_i(\theta)$. If the $\phi_i(\theta)$ are linear this implies that both numerator and denominator of $S^0(\theta)$ and $S^*(\theta)$ are convex functions of θ. Consequently, they are continuous and differentiable, except possibly at a finite number of points. By using a general result of Rockafeller (1970, corollary 26-3-2 and the remarks following it) on the differentiability of the mappings

$$S(x) = \inf_{y \varepsilon C} |x-y|^p$$

with $p > 1$ and C an arbitrary closed nonempty convex set, we find that the numerator of $S^0(\theta)$ is continuously differentiable. Consequently, $S^0(\theta)$ is continuously differentiable at all points where the denominator does not vanish. A more direct proof of this result has been given by Kruskal (1971). In general, the second partials of $S^0(\theta)$ and $S^*(\theta)$ need not exist, even if the $\phi_i(\theta)$ are infinitely differentiable. For computational purposes this is, of course, not very important, but it certainly affects the applicability of existing convergence proofs for (modified) gradient methods.

REFERENCES

Guttman, L. A general nonmetric technique for finding the smallest coordinate space for a configuration of points. Psychometrika, 1968, 33, 469-506.

Kruskal, J. B. Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. Psychometrika, 1964, 29, 1-28. (a)

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. Psychometrika, 1964, 29, 115-129. (b)

Kruskal, J. B. Analysis of factorial experiments by estimating monotone transformation of the data. J. Roy. Statist. Soc., B, 1965, 27, 251-263.

Kruskal, J. B. Monotone regression: Continuity and differentiability properties. Psychometrika, 1971, 36, 57-62.

Rockafellar, R. T. Convex analysis. Princeton, N.J.: Princeton Univ. Press, 1970.

No:	Region No.	Rank Image Y^*	Block Almagamation Y^o	$S^*(\theta)$	$S^o(\theta)$
1:	$0 < \theta_1 < \theta_2$	$Y_1 < Y_2 < Y_3 < Y_4$	$Y_{[1]} < Y_{[2]} < Y_{[3]} < Y_{[4]}$	0	0
2:	$0 < \theta_1 = \theta_2$	$Y_1 < Y_2 = Y_3 < Y_4$	$Y_{[1]} < Y_{[2]} = Y_{[3]} < Y_{[4]}$	0	0
3:	$0 < \theta_2 < \theta_1$	$Y_1 < Y_3 < Y_2 < Y_4$	$Y_1 < Y_{[2,3]} = Y_{[2,3]} < Y_4$	$\frac{1}{2}(\theta_1-\theta_2)^2$	$\frac{1}{2}(\theta_1-\theta_2)^2$
4:	$0 = \theta_2 < \theta_1$	$Y_1 = Y_3 < Y_2 = Y_4$	$Y_1 < Y_{[2,3]} = Y_{[2,3]} < Y_4$	$\frac{1}{2}$	$\frac{1}{2}$
5:	$0 < -\theta_2 < \theta_1$	$Y_3 < Y_1 < Y_4 < Y_2$	$Y_1 < Y_{[2,3]} = Y_{[2,3]} < Y_4$	$\frac{1}{2}$	$\frac{1}{2}(\theta_1-\theta_2)^2$
6:	$0 < -\theta_2 = \theta_1$	$Y_3 < Y_1 = Y_4 < Y_2$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}$	1
7:	$0 < \theta_1 < -\theta_2$	$Y_3 < Y_4 < Y_1 < Y_2$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	θ_2^2	1
8:	$0 = \theta_1 < -\theta_2$	$Y_3 = Y_4 < Y_1 = Y_2$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	1	1
9:	$\theta_2 < \theta_1 < 0$	$Y_4 < Y_3 < Y_2 < Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	1	1
10:	$\theta^2 = \theta_2 < 0$	$Y_4 < Y_3 = Y_2 < Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	1	1
11:	$\theta_1 < \theta_2 < 0$	$Y_4 < Y_2 < Y_3 < Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}(\theta_1+\theta_2)^2$	1
12:	$\theta_1 < \theta_2 = 0$	$Y_4 = Y_2 < Y_3 = Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}$	1
13:	$\theta_1 < -\theta_2 < 0$	$Y_2 < Y_4 < Y_1 < Y_3$	$Y_{[1,2]} = Y_{[1,2]} < Y_{[3,4]} = Y_{[3,4]}$	$\frac{1}{2}$	θ_1^2
14:	$\theta_1 = -\theta_2 < 0$	$Y_2 < Y_4 = Y_1 < Y_3$	$Y_{[1,2]} = Y_{[1,2]} < Y_{[3,4]} = Y_{[3,4]}$	$\frac{1}{2}$	$\frac{1}{2}$
15:	$-\theta_2 < \theta_1 < 0$	$Y_2 < Y_1 < Y_4 < Y_3$	$Y_{[1,2]} = Y_{[1,2]} < Y_{[3,4]} = Y_{[3,4]}$	θ_1^2	θ_1^2
16:	$-\theta_2 < \theta_1 = 0$	$Y_2 = Y_1 < Y_4 = Y_3$	$Y_{[1]} = Y_{[2]} < Y_{[3]} = Y_{[4]}$	0	0

Table 1:1

No	Region	Y^*	Y°	$S^*(\theta)$	$S^\circ(\theta)$
1	$0 \le \theta_1 \le \theta_2$	$Y_1 \le Y_4 \le Y_3 \le Y_2$	$Y_{[1]} < Y_{[2,3,4]} = Y_{[2,3,4]} = Y_{[2,3,4]}$	$\frac{1}{2}\theta_2^2$	$\frac{2}{3}(1-\theta_1\theta_2)$
2a	$\theta_2 \le \theta_1 \le 2\theta_2$	$Y_1 \le Y_3 \le Y_4 \le Y_2$	$Y_{[1]} < Y_{[2,3,4]} = Y_{[2,3,4]} = Y_{[2,3,4]}$	$\frac{1}{2}[1-\theta_1\theta_2]$	$\frac{2}{3}(1-\theta_1\theta_2)$
2b	$0 \le \theta_2 \le \frac{1}{2}\theta_1$	$Y_1 \le Y_3 \le Y_4 \le Y_2$	$Y_{[1]} < Y_{[2,3]} = Y_{[2,3]} < Y_{[4]}$	$\frac{1}{2}[1-\theta_1\theta_2]$	$\frac{1}{2}\theta_1^2$
3a	$\frac{1}{2}\theta_1 \le \theta_2 \le 0$	$Y_3 \le Y_1 \le Y_2 \le Y_4$	$Y_{[1]} < Y_{[2,3]} = Y_{[2,3]} < Y_{[4]}$	$\frac{1}{2}[1+\theta_1\theta_2]$	$\frac{1}{2}\theta_1^2$
3b	$2\theta_2 \le -\theta_1 \le \theta_2$	$Y_3 \le Y_1 \le Y_2 \le Y_4$	$Y_{[1,2,3]} = Y_{[1,2,3]} = Y_{[1,2,3]} < Y_{[4]}$	$\frac{1}{2}[1+\theta_1\theta_2]$	$\frac{2}{3}(1+\theta_1\theta_2)$
4	$\theta_2 \le -\theta_1 \le 0$	$Y_3 \le Y_2 \le Y_1 \le Y_4$	$Y_{[1,2,3]} = Y_{[1,2,3]} = Y_{[1,2,3]} < Y_{[4]}$	$\frac{1}{2}\theta_2^2$	$\frac{2}{3}(1+\theta_1\theta_2)$
5	$\theta_2 \le \theta_1 \le 0$	$Y_2 \le Y_3 \le Y_4 \le Y_1$	$Y_{[1,2,3]} = Y_{[1,2,3]} = Y_{[1,2,3]} < Y_{[4]}$	$\theta_1^2 + \frac{1}{2}\theta_2^2$	$\frac{2}{3}(1+\theta_1\theta_2)$
6	$\theta_1 \le \theta_2 \le 0$	$Y_2 \le Y_4 \le Y_3 \le Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}(1+\theta_1\theta_2)$	1
7	$\theta_1 \le -\theta_1 \le 0$	$Y_4 \le Y_2 \le Y_1 \le Y_3$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}(1-\theta_1\theta_2)$	1
8	$-\theta_2 \le \theta_1 \le 0$	$Y_4 \le Y_1 \le Y_3 \le Y_2$	$Y_{[1]} < Y_{[2,3,4]} = Y_{[2,3,4]} = Y_{[2,3,4]}$	$\theta_1^2 + \frac{1}{2}\theta_2^2$	$\frac{2}{3}(1-\theta_1\theta_2)$

Table 2:1

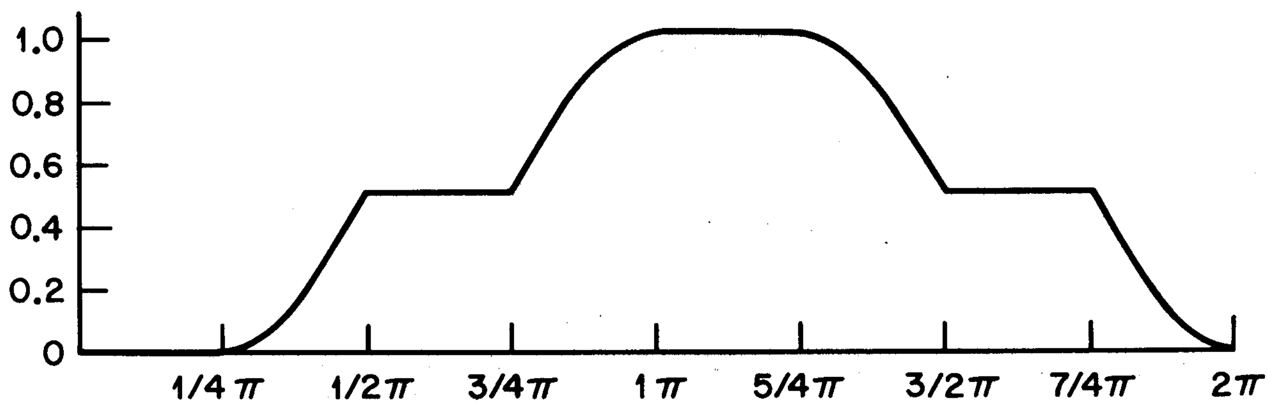


Fig. 1:2

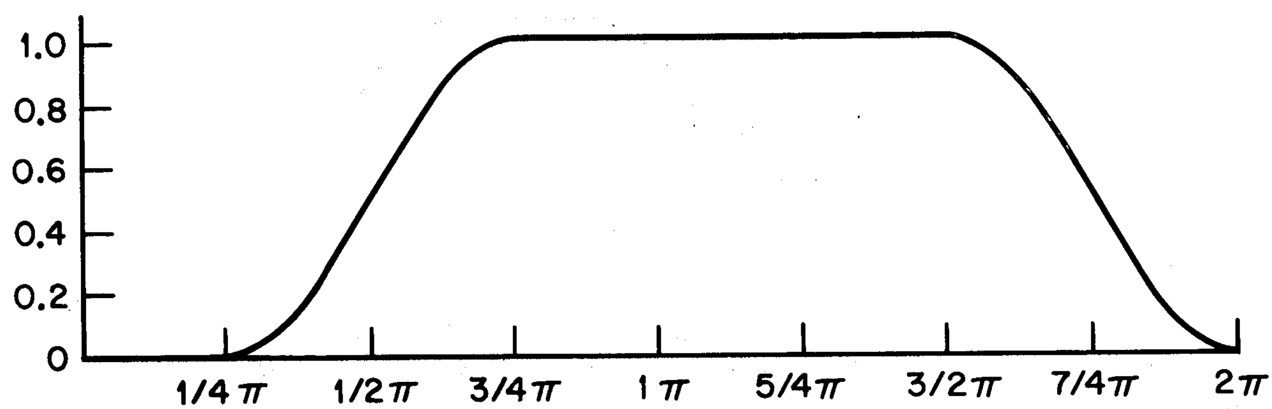


Fig. 1:3

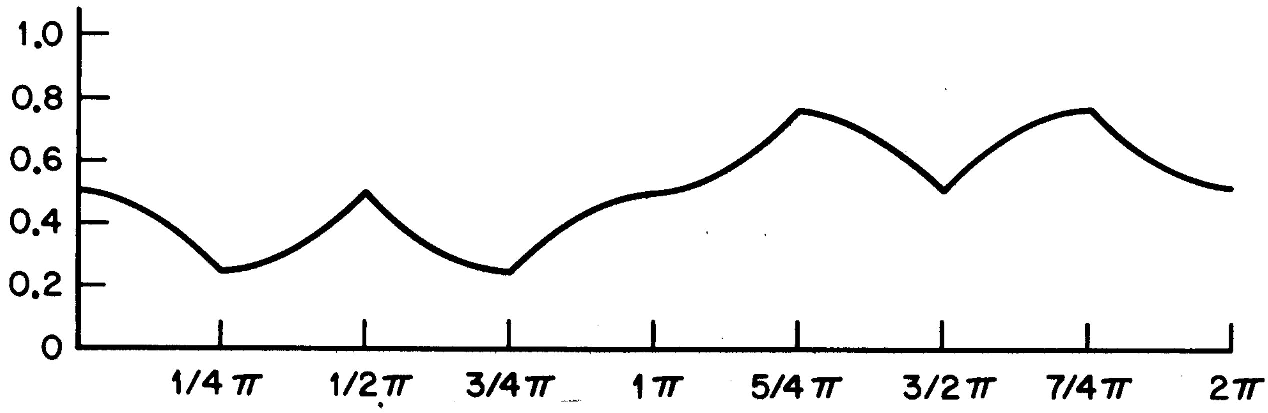


Fig. 2:2

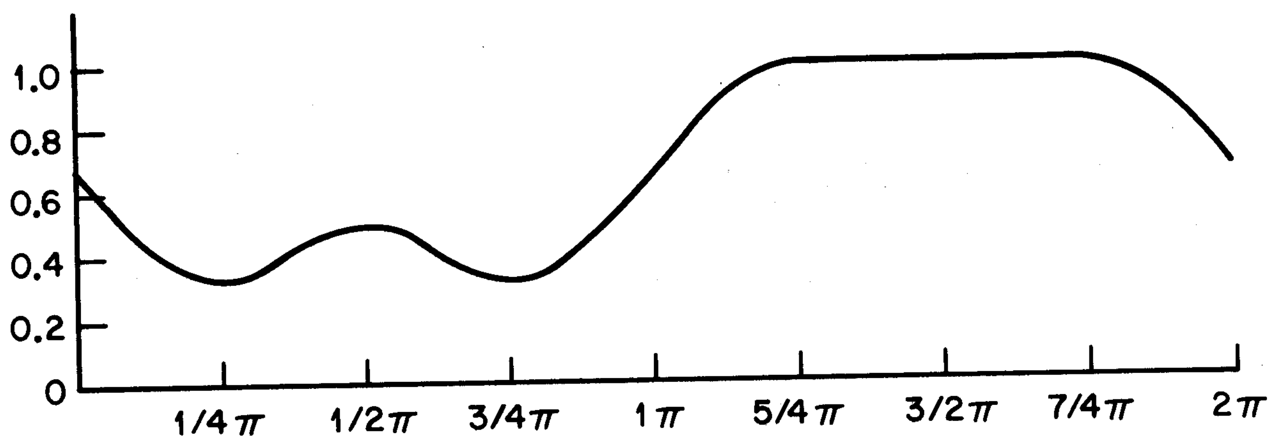


Fig. 2:3

Comments on

[redacted] SMOOTHNESS PROPERTIES OF NONMETRIC LOSS FUNCTIONS

by JAN DELEEU[redacted]W

for the author. June 20, 1974

Joseph B Kruskal

[redacted]

In two key places, the author fails to explain what [redacted] the central [redacted] point of importance.

The first of these is really vital, since it coneerns the central point of the whole paper. In the introduction, the central point concerns the nonexistence of the derivatives at the desired optimum [redacted] configurations. Nonexistence of the derivatives on a set of measure 0 could be of very little importance to a practical numerical method. Nonexistence of the derivative at the desired solution, on the contrary, has very strong practical implications. In particular, it means that the most common approach to optimization, which involves finding a place where the gradient is zero, will at best work poorly, and [redacted] depending on the circumstances may not work at all.

Associated with this failure of explanation is a much more minor one, which is however still important. [redacted] Nonexistence of derivatives is something which can occur in many ways. [redacted] Where possible, it is [redacted] far better to refer specifically to the manner in which the derivative fails to exist. This is indeed possible here. [redacted]

Combining this with the preceding, [redacted] I would urge an introduction which starts something like this:

[redacted] "Most of the computational approaches [redacted] in the nonmetric scaling area completely ignore the possibility that the [redacted] optimum [redacted] may occur [redacted] where the loss function [redacted] is not smooth, such as at a sharp ridge or valley. [redacted] This possibility is of substantial practical importance, since optima which occur at such points [redacted] cannot be [redacted] efficiently located by the most common approach to optimization, namely, finding the zero of the gradient function, [redacted] and in unfavorable cases [redacted] may not be located all. [redacted]

¶ We shall show that [redacted]s Guttman's rank-image method [redacted] __usually__ [redacted] has its optimum value at such a sharp ridge, and that [redacted] this difficulty is [redacted] intrinsic [redacted] in the rank-image method. This explains certain oscillations which have been [redacted] been observed in the Guttman-Lingoes SSA programs.

Our general approach, [redacted] which rests on some powerful theorems, also yields much more simply the theorem of [redacted] Kruskal(1971) which states that the same phenonenon does __not__ occur in his monotonic regression approach, and permits a similar approach to other loss functions in the future."

The second key place where De Leeuw fails to make his central point clear occurs on page 2, starting at [redacted] the middle. [redacted] This failure [redacted] contains several components.

(a) He fails to make clear [redacted] how the [redacted] model function A( ) and the [redacted] loss function are related.

(b) He fails to make clear that he is introducing the Kruskal and Guttman loss functions at this point as illustrations of how his [redacted] terminology [redacted] works.

(c) He says which properties of the loss function [redacted] he is <u>not</u> concerned with, but does not say which ones he <u>is</u> concerned with.

(d) etc.

Based on all this, I would suggest that the section labelled "Theoretical Results" start something like this:

"In this paper we consider loss functions of the form

[redacted]

$$S(\theta) = N[A(\theta) - T(A(\theta))] / N[A(\theta)].$$

[redacted]

Here N is the usual norm [redacted] consisting of the sum of squares of the vector. The mapping $y = A(\theta)$ [redacted] defines the model in question. For example, in the case of multidimensional scaling, A is the mapping from [redacted] a configuration of points to the vector of all interpoint distances. In general, $A:\Theta \to Y$ is a [redacted] (not necessarily linear) vector-valued mapping defined on some open subset Θ of Euclidean space into the space of all y vectors. Finally, $T:Y \to Y$ is a [redacted] mapping which depends on the data, and [redacted] embodies the distinction between the Kruskal monotonic regression approach, the Guttman [redacted] rank-image approach, and other [redacted] related approaches. [redacted]

[redacted] We define $T^0(y)$ to be the vector y^0 which [redacted] satisfies the rank order restrictions of the data (that is, which is monotonic in the observations) and which [redacted] minimizes [redacted] $N(y-y^0)$ [redacted]. We define $T^*(y)$ to be the vector y^* which is [redacted] the permutation of the elements of y which satisfies the order restrictions in the data and which minimizes $N(y - y^*)$. [redacted]

[redacted] We let S^0 [redacted] and S^* be the S functions using T^0 and T^*. [redacted]

(This is a slight generalization of Guttman's rank-image idea, which makes sense for partial orders as well as complete [redacted] orders.)

[redacted]

Then S^0 is the square of Kruskal's [redacted] stress function, and S^* is [redacted] is related to Guttman's coefficient of alienation K by a simple relationship, namely, $K = \sqrt{1 - S^*}$.

[redacted]

There [redacted] is a host of results [redacted] in the mathematical literature on the properties of the maps [redacted] $y \to T^0(y)$ and $y \to N[y - T^0(y)]$. We mention ..... [redacted]

. . . . . .

and that [redacted] $y^0(A(\theta))$ is a uniformly continuous, almost everywhere differentiable [redacted] function of [redacted] θ.

[redacted]

Some [redacted] of these results have parallels [redacted] for T^* and S^*, though the material is not so well-known. [redacted] However, as the examples in the next section will show, one crucial result does not hold, namely, the derivatives of S^* fail to exist at a number of points, even if A is linear. Furthermore, [redacted] when A is linear, the derivatives of S^* always fail to [redacted] exist at the [redacted] optimum (that is, the minimum) position for S^*. When [redacted] A is non-linear, it is difficult [redacted] to state any precise theorem, but it would appear that the derivatives [redacted] "usually" fail to exist at the optimum. What does hold for T^* and S^* is this: ...... "

Red is more important

SMOOTHNESS PROPERTIES OF NONMETRIC LOSS FUNCTIONS

Jan DeLeeuw*
Bell Laboratories
Murray Hill, New Jersey 07974

0: Introduction

One of the more spectacular ingredients of most papers ~~on~~ in the nonmetric scaling area is a long and complicated formula which purports to give the partial derivatives of the loss function with respect to the parameters. The question of existence of these derivatives is usually ignored. A notable exception is the paper by Kruskal (1971). The general impression, however, is that most authors think that the nonexistence of the derivatives at a number of points is of little practical consequence. In a sense this is true. The computer programs are not bothered by these difficulties during their iterations, and for some of the more important approaches to scaling the results of Kruskal guarantee sufficient differentiability. In this paper we shall show that the results are important for the practical implementation of Guttman's rank-image method. We first derive some results on differentiability and review some of the relevant literature. The results are illustrated by some examples, which are also of some independent interest. In fact they are the first examples which show the complete behavior of nonmetric loss functions

---

*On leave of absence from the University of Leiden, Leiden, The Netherlands.

over the whole parameter space, and they clearly illustrate the piecewise smoothness of these loss functions. The main technical result is that Kruskal's monotone regression method joins the pieces in a smooth way, while Guttman's rank image method joins them in a nonsmooth way.

1: Theoretical Results

In this paper we are not concerned with the properties of the loss function that are due entirely to properties of the model (or the combination rule). It is clear that for some of the models discussed by Young (1972) the differentiability properties of the loss function are destroyed by the nondifferentiability of the combination rules. In our development we shall use a completely general model of the form $y = A(\theta)$ with A a general (not necessarily linear) vector valued mapping defined on an open subset Θ of a Euclidean space. We define the following versions of the Kruskal (1964 a,b, 1965) and Guttman (1968) loss functions:

prime

$$S^0 \triangleq \frac{(y-y^0)^1(y-y^0)}{y^1 y}$$

$$S^* \triangleq \frac{(y-y^*)^1(y-y^*)}{y^1 y}$$

Here y^0 is defined as the vector that satisfied the order restrictions in the data, and minimizes S^0 for fixed y. The vector y* is defined as that permutation of the elements of y that satisfies the order restrictions in the data, and minimizes S* for fixed y. This last definition is a slight generalization of Guttman's idea, which also makes sense for partial orders. There is a host of results in the mathematical literature on the properties of the maps $y \rightarrow y^0$ and $y \rightarrow N^0$. We mention Moreau (1965), Kruskal (1969, 1971), Rockafellar (1970, p. 255), Asplund (1968, 1973), Holmes (1973). Using these results and assuming

prime

that $y'y$ does not vanish on Θ and that $A(\theta)$ is sufficiently many times differentiable, it follows easily that S^0 is a continuously differentiable, almost everywhere twice differentiable, function of θ and that y^0 is a uniformly continuous, almost everywhere differentiable, function of θ. For S* and y* the results are less well known. It remains true, however, that $||y_1^*-y_2^*|| \leq ||y_1-y_2||$, which implies that y* is an almost very differentiable function of y. Together with the usual assumptions on $A\theta$ this makes S* an almost everywhere differentiable function of θ. The examples in the next section will show that the derivatives of S* actually fail to exist at a number of points, even if $A\theta$ is linear.

2: A First Example

Consider the linear model $y = A(\theta)$ given by

$$y = \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} -\theta_1-\theta_2 \\ \theta_1-\theta_2 \\ -\theta_1+\theta_2 \\ \theta_1+\theta_2 \end{bmatrix}$$

~~We desire~~ The corresponding data values x_i satisfy $x_1 < x_2 < x_3 < x_4$, so that we want the Y_i to be as nearly as possible weakly ascending in i.

$$\cancel{Y_1 \le Y_2 \le Y_3 \le Y_4}.$$

Without loss of generality we restrict ourselves to values of θ on the unit circle,

$$\theta_1^2 + \theta_2^2 = 1 \quad .$$

We define

$$S^*(\theta) = \frac{1}{16}\{(Y_1-Y_1^*)^2 + (Y_2-Y_2^*)^2 + (Y_3-Y_3^*)^2 + (Y_4-Y_4^*)^2\},$$

with ~~Y^*~~ $y^* = T^*(y)$ a permutation of ~~Y~~ y satisfying

$$Y_1^* \le Y_2^* \le Y_3^* \le Y_4^*.$$

We also define

$$S^0(\theta) = \tfrac{1}{4}\{(Y_1-Y_1^0)^2 + (Y_2-Y_2^0)^2 + (Y_3-Y_3^0)^2 + (Y_4-Y_4^0)^2\},$$

with ~~Y^0~~ the (unique) vector satisfying

$$Y_1^0 \leq Y_2^0 \leq Y_3^0 \leq Y_4^0$$

and minimizing ~~S^0~~ for fixed Y. Both coefficients ~~are~~ rescaled in such a way that they vary in the closed interval [0,1].

Moreover,

$$
\begin{aligned}
S^0(\theta) = 0 \quad \text{iff } S^*(\theta) = 0 \quad & \text{iff } Y^0 = Y^* = Y. \\
& \text{iff } Y_1 \leq Y_2 \leq Y_3 \leq Y_4. \\
& \text{iff } 0 \leq \theta_1 \leq \theta_2
\end{aligned}
$$

and

$$
\begin{aligned}
S^*(\theta) = 1 \quad \text{iff } Y^* = -Y \quad & \text{iff } Y_4 \leq Y_3 \leq Y_2 \leq Y_1 \\
& \text{iff } \theta_2 \leq \theta_1 \leq 0
\end{aligned}
$$

and

$$
\begin{aligned}
S^0(\theta) = 1 \quad \text{iff } Y^0 = 0 \quad & \text{iff } \{ Y_{[2,3,4]} \leq Y_{[1]} \\
& \text{or } Y_{[3,4]} \leq Y_{[1,2]} \text{ and } Y_{[4]} \leq Y_{[1,2,3]} \} \\
& \text{iff } \theta_2 \leq 0 \text{ and } \theta_1 + \theta_2 \leq 0
\end{aligned}
$$

& becomes && (with spaces)

We use the notation $Y_{[i_1,\dots,i_p]}$ for the average of $Y_{i_1},\dots,Y_{i_p}$. Table 1:1 gives the different regions ~~(going around the circle)~~ in (θ_1,θ_2) space with their different characterizations in cyclic order. Figure 1:2 gives a plot of $S^*(\theta)$ using ~~polar~~ angular coordinates, and 1:3 gives a similar plot of $S^0(\theta)$. From the tables and figures it is clear that both functions are continuous. Moreover, $S^0(\theta)$ is continuously differentiable, while the derivatives of $S^*(\theta)$ do not exist at the points $\frac{1}{2}\pi$, $\frac{3}{4}\pi$, $\frac{3}{2}\pi$, and $\frac{7}{4}\pi$. The second derivatives of $S^*(\theta)$ do not exist at all eight "special" points and those of $S^0(\theta)$ do not exist at $\frac{1}{4}\pi$, $\frac{3}{4}\pi$, $\frac{3}{2}\pi$, and 2π. The concave appearance of the function is deceptive; if we had broken open the circle at π, then the one-dimensional plot would have been convex-like. Of course, this example is unrealistic for various reasons. In the first place, a perfect solution exists. In the second place, there are no real local minima except the platform of $S^*(\theta)$ between $\frac{1}{2}\pi$ and $\frac{3}{4}\pi$ and between $\frac{3}{4}\pi$ and $\frac{7}{4}\pi$. In real examples, perfectly flat areas like this usually will not occur. There may, of course, be a considerable region with $S^*(\theta) = 1$ and/or $S^0(\theta) = 1$.

3: A Second Example

It is clear from Table 1:1 that the linear model

$$y = \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} -\theta_1-\theta_2 \\ \theta_1+\theta_2 \\ \theta_1+\theta_2 \\ \theta_1-\theta_2 \end{bmatrix}$$

with requirements

$$Y_1 \leq Y_2 \leq Y_3 \leq Y_4$$

does not have a perfect solution. The relevant information is contained in Table 2:1, and the plots are drawn in 2:2 and 2:3.

The results are quite different in several respects. In the first place, there are a number of stationary values. The function $S^*(\theta)$ has local maxima at $\frac{1}{2}\pi$, $\frac{5}{4}\pi$, $\frac{7}{4}\pi$, local minima at $\frac{1}{4}\pi$, $\frac{3}{4}\pi$, $\frac{3}{4}\pi$, and saddle values at π, 2π. The function is not differentiable at all these minima and maxima. The function $S^0(\theta)$ is more smooth. It has a local maximum at $\frac{1}{2}\pi$, and two local minima at $\frac{1}{4}\pi$ and $\frac{3}{4}\pi$ (both local minima are also global). It is continuously differentiable at all points, but at the special points the second order partials do not exist.

4: Practical Implications

The most important practical consequence of our result is that it seems to be quite common that the derivatives of S^* do not exist at local minima, and, more seriously, are bounded ~~array~~ away from zero in ~~each~~ the neighborhood of those minima. This implies that ordinary gradient methods cannot work very well in the minimization of S^*; it also means that ~~the purpose of the~~ rank image method cannot be formulated as finding a stationary point of S*. Fortunately there exists an extensive literature on alternative necessary conditions for a local minimum and alternat~~e~~ive minimization algorithms for functions which are merely directionally differentiable,

Though the existing methods are a good deal more complicated and expensive ~~that~~ than the methods for finding a stationary point of S*

REFERENCES

Asplund, E. Fréchet differentiability of convex mappings. Acta Math., 1968, 121, 31-48.

Asplund, E. Differentiability of the metric projection in finite dimensional Euclidean space. *Proc. A.M.S.*, 1973, 38, 213-219.

Guttman, L. A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, 1968, 33, 469-506.

Holmes, R. B. Smoothness of certain metric projections on Hilbert space. *Trans. A.M.S.*, 1973, 184, 87-100.

Kruskal, J. B. Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. *Psychometrika*, 1962, 29, 1-28. (a)

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 1964, 29, 115-129. (b)

Kruskal, J. B. Analysis of factorial experiments by estimating monotone transformation of the data. *J. Roy. Statist. Soc.*, B, 1965, 27, 251-263.

Kruskal, J. B. Two convex counter examples: A discontinuous envelope function and a nondifferentiable nearest point mapping. *Proc. A.M.S.*, 1967, 23, 697-703.

Kruskal, J. B. Monotone regression: Continuity and differentiability properties. *Psychometrika*, 1971, 36, 57-62.

Moreau, J. J. Proximité et dualité dans un espace Hilbertien. Bull. Soc. Math. France, 1965, 93, 273-299.

Moreau, J. J. Inf-convolution, sous-additivité, convexité des functions numeriques. J. Math. Pures et appl., 1970, 49, 109-154.

Rockafellar, R. T. Convex analysis. Princeton, N. J.: Princeton Univ. Press, 1970.

Young, F. W. A model for polynomial conjoint analysis algorithms. In R. N. Shepard, A. K. Romney, S. B. Nerlove (Eds.), Multidimensional scaling: Theory and applications in the behavioral sciences. New York: Seminar Press, 1972. Pp. 69-104.

No:	Region No.	Rank Image Y^*	Block Amalgamation Y^o	$S^*(\theta)$	$S^o(\theta)$
1:	$0 < \theta_1 < \theta_2$	$Y_1 < Y_2 < Y_3 < Y_4$	$Y_{[1]} < Y_{[2]} < Y_{[3]} < Y_{[4]}$	0	0
2:	$0 < \theta_1 = \theta_2$	$Y_1 < Y_2 = Y_3 < Y_4$	$Y_{[1]} < Y_{[2]} = Y_{[3]} < Y_{[4]}$	0	0
3:	$0 < \theta_2 < \theta_1$	$Y_1 < Y_3 < Y_2 < Y_4$	$Y_1 < Y_{[2,3]} = Y_{[2,3]} < Y_4$	$\frac{1}{2}(\theta_1-\theta_2)^2$	$\frac{1}{2}(\theta_1-\theta_2)^2$
4:	$0 = \theta_2 < \theta_1$	$Y_1 = Y_3 < Y_2 = Y_4$	$Y_1 < Y_{[2,3]} = Y_{[2,3]} < Y_4$	$\frac{1}{2}$	$\frac{1}{2}$
5:	$0 < -\theta_2 < \theta_1$	$Y_3 < Y_1 < Y_4 < Y_2$	$Y_1 < Y_{[2,3]} = Y_{[2,3]} < Y_4$	$\frac{1}{2}$	$\frac{1}{2}(\theta_1-\theta_2)^2$
6:	$0 < -\theta_2 = \theta_1$	$Y_3 < Y_1 = Y_4 < Y_2$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}$	1
7:	$0 < \theta_1 < -\theta_2$	$Y_3 < Y_4 < Y_1 < Y_2$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	θ_2^2	1
8:	$0 = \theta_1 < -\theta_2$	$Y_3 = Y_4 < Y_1 = Y_2$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	1	1
9:	$\theta_2 < \theta_1 < 0$	$Y_4 < Y_3 < Y_2 < Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	1	1
10:	$\theta^2 = \theta_2 < 0$	$Y_4 < Y_3 = Y_2 < Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	1	1
11:	$\theta_1 < \theta_2 < 0$	$Y_4 < Y_2 < Y_3 < Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}(\theta_1+\theta_2)^2$	1
12:	$\theta_1 < \theta_2 = 0$	$Y_4 = Y_2 < Y_3 = Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}$	1
13:	$\theta_1 < -\theta_2 < 0$	$Y_2 \quad Y_4 \quad Y_1 \quad Y_3$	$Y_{[1,2]} = Y_{[1,2]} < Y_{[3,4]} = Y_{[3,4]}$	$\frac{1}{2}$	θ_1^2
14:	$\theta_1 = -\theta_2 < 0$	$Y_2 < Y_4 = Y_1 < Y_3$	$Y_{[1,2]} = Y_{[1,2]} < Y_{[3,4]} = Y_{[3,4]}$	$\frac{1}{2}$	$\frac{1}{2}$
15:	$-\theta_2 < \theta_1 < 0$	$Y_2 < Y_1 < Y_4 < Y_3$	$Y_{[1,2]} = Y_{[1,2]} < Y_{[3,4]} = Y_{[3,4]}$	θ_1^2	θ_1^2
16:	$-\theta_2 < \theta_1 = 0$	$Y_2 = Y_1 < Y_4 = Y_3$	$Y_{[1]} = Y_{[2]} < Y_{[3]} = Y_{[4]}$	0	0

Table 1:1

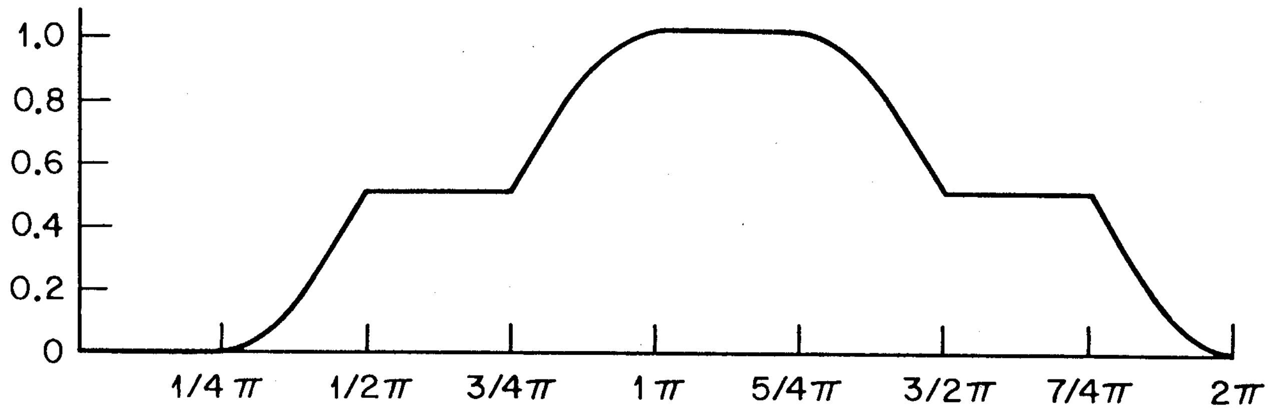


Fig. 1:2

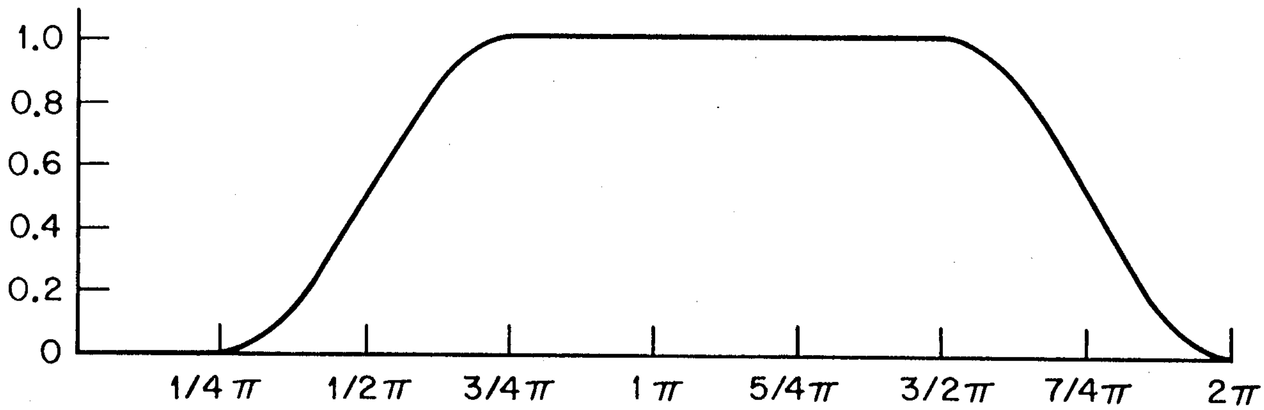


Fig. 1:3

No	Region	Y^*	Y^o	$S^*(\theta)$	$S^o(\theta)$
1	$0 \le \theta_1 \le \theta_2$	$Y_1 \le Y_4 \le Y_3 \le Y_2$	$Y_{[1]} < Y_{[2,3,4]} = Y_{[2,3,4]} = Y_{[2,3,4]}$	$\frac{1}{2}\theta_2^2$	$\frac{2}{3}(1-\theta_1\theta_2)$
2a	$\theta_2 \le \theta_1 \le 2\theta_2$	$Y_1 \le Y_3 \le Y_4 \le Y_2$	$Y_{[1]} < Y_{[2,3,4]} = Y_{[2,3,4]} = Y_{[2,3,4]}$	$\frac{1}{2}[1-\theta_1\theta_2]$	$\frac{2}{3}(1-\theta_1\theta_2)$
2b	$0 \le \theta_2 \le \frac{1}{2}\theta_1$	$Y_1 \le Y_3 \le Y_4 \le Y_2$	$Y_{[1]} < Y_{[2,3]} = Y_{[2,3]} < Y_{[4]}$	$\frac{1}{2}[1-\theta_1\theta_2]$	$\frac{1}{2}\theta_1^2$
3a	$\frac{1}{2}\theta \le \theta_2 \le 0$	$Y_3 \le Y_1 \le Y_2 \le Y_4$	$Y_{[1]} < Y_{[2,3]} = Y_{[2,3]} < Y_{[4]}$	$\frac{1}{2}[1+\theta_1\theta_2]$	$\frac{1}{2}\theta_1^2$
3b	$2\theta_2 \le -\theta_1 \le \theta_2$	$Y_3 \le Y_1 \le Y_2 \le Y_4$	$Y_{[1,2,3]} = Y_{[1,2,3]} = Y_{[1,2,3]} < Y_{[4]}$	$\frac{1}{2}[1+\theta_1\theta_2]$	$\frac{2}{3}(1+\theta_1\theta_2)$
4	$\theta_2 \le -\theta_1 \le 0$	$Y_3 \le Y_2 \le Y_1 \le Y_4$	$Y_{[1,2,3]} = Y_{[1,2,3]} = Y_{[1,2,3]} < Y_{[4]}$	$\frac{1}{2}\theta_2^2$	$\frac{2}{3}(1+\theta_1\theta_2)$
5	$\theta_2 \le \theta_1 \le 0$	$Y_2 \le Y_3 \le Y_4 \le Y_1$	$Y_{[1,2,3]} = Y_{[1,2,3]} = Y_{[1,2,3]} < Y_{[4]}$	$\theta_1^2 + \frac{1}{2}\theta_2^2$	$\frac{2}{3}(1+\theta_1\theta_2)$
6	$\theta_1 \le \theta_2 \le 0$	$Y_2 \le Y_4 \le Y_3 \le Y_1$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}(1+\theta_1\theta_2)$	1
7	$\theta_1 \le -\theta_1 \le 0$	$Y_4 \le Y_2 \le Y_1 \le Y_3$	$Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]} = Y_{[1,2,3,4]}$	$\frac{1}{2}(1-\theta_1\theta_2)$	1
8	$-\theta_2 \le \theta_1 \le 0$	$Y_4 \le Y_1 \le Y_3 \le Y_2$	$Y_{[1]} < Y_{[2,3,4]} = Y_{[2,3,4]} = Y_{[2,3,4]}$	$\theta_1^2 + \frac{1}{2}\theta_2^2$	$\frac{2}{3}(1-\theta_1\theta_2)$

Table 2:1

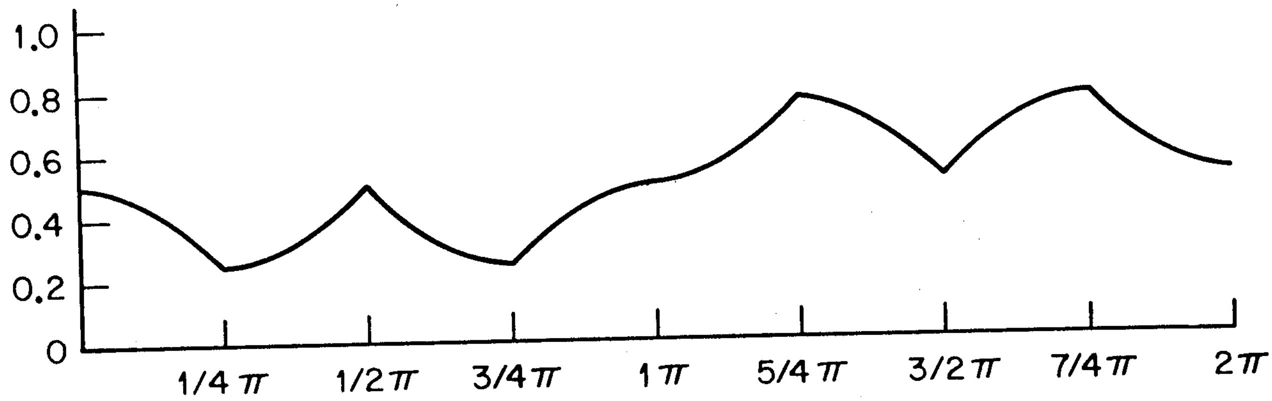


Fig. 2:2

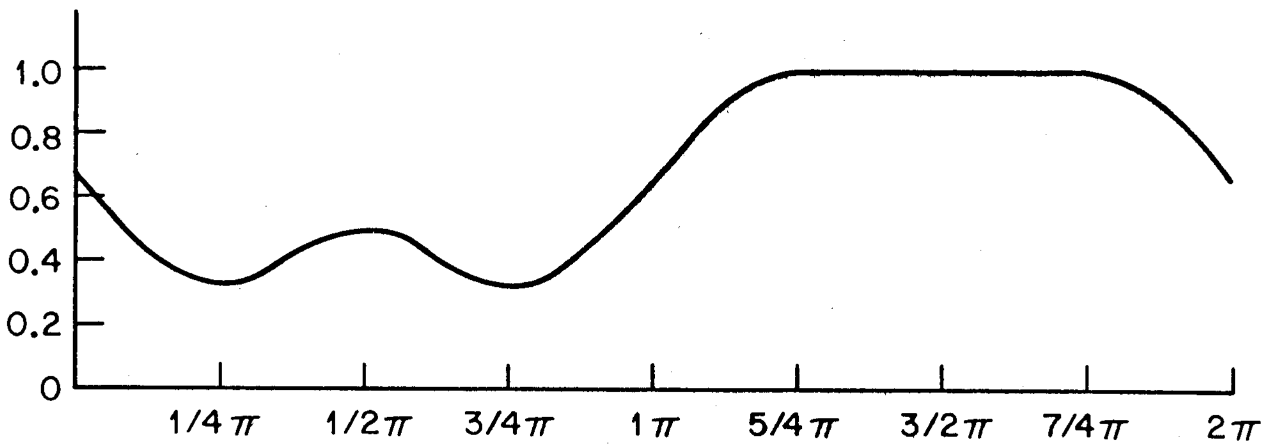


Fig. 2:3

Bell Laboratories

600 Mountain Avenue
Murray Hill, New Jersey 07974
Phone (201) 582-3000

March 27, 1974

Professor L. Guttman
Israel Institute of Applied
Social Research
George Washington 19
P.O.B. 7150
Jerusalem
ISRAEL

Dear Louis,

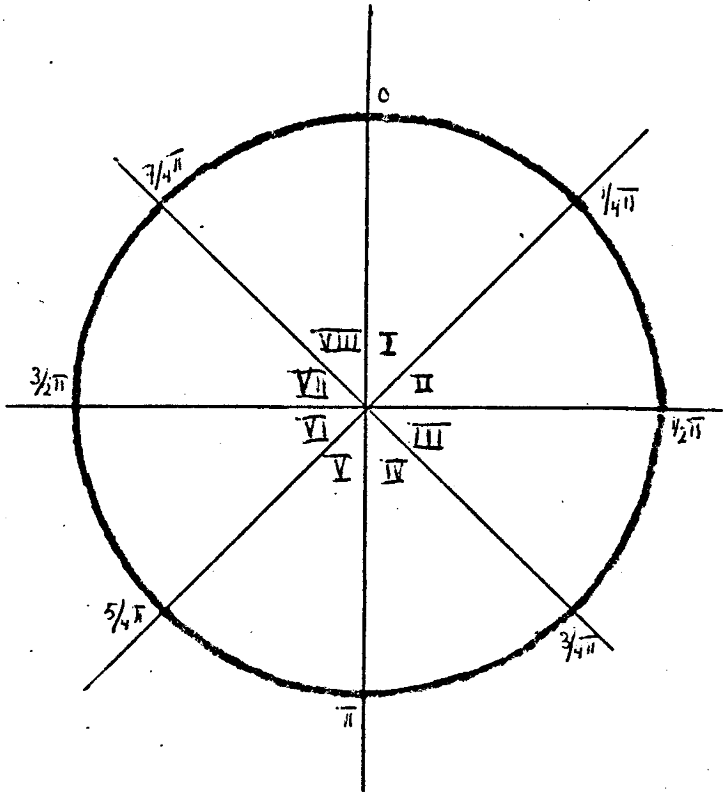
I recently came across a difficulty in the rank-image approach which seems to be quite fundamental from a theoretical point of view, although I am not altogether sure what the precise practical consequences are. Consider the linear model $y = A\theta$ with A equal to

$$\begin{pmatrix} -1 & -1 \\ +1 & +1 \\ -1 & +1 \\ +1 & -1 \end{pmatrix}$$

and with ordinal restrictions $y_1 \leq y_2 \leq y_3 \leq y_4$. We study the behavior of the rank-image loss function

$$S^* = \frac{1}{16} (y-y^*)'(y-y^*)$$

on the circle $\theta'\theta = 1$. The factor $\frac{1}{16}$ is chosen so that $0 \leq S^* \leq 1$. The rank-image transformation differs, of course, from one region to another. We have to distinguish eight different cases



Region	Order Relations	S^*	$dS^*/d\xi$
1	$y_1 \le y_4 \le y_3 \le y_2$	$\frac{1}{2}\cos^2\xi$	$-\cos\xi\sin\xi$
2	$y_1 \le y_3 \le y_4 \le y_2$	$\frac{1}{2}[1-\sin\xi\cos\xi]$	$\frac{1}{2}[\sin^2\xi-\cos^2\xi]$
3	$y_3 \le y_1 \le y_2 \le y_4$	$\frac{1}{2}[1+\sin\xi\cos\xi]$	$\frac{1}{2}[\cos^2\xi-\sin^2\xi]$
4	$y_3 \le y_2 \le y_1 \le y_4$	$\frac{1}{2}\cos^2\xi$	$-\cos\xi\sin\xi$
5	$y_2 \le y_3 \le y_4 \le y_1$	$\sin^2\xi + \frac{1}{2}\cos^2\xi$	$\cos\xi\sin\xi$
6	$y_2 \le y_4 \le y_3 \le y_1$	$\frac{1}{2}[1+\sin\xi\cos\xi]$	$\frac{1}{2}[\cos^2\xi-\sin^2\xi]$
7	$y_4 \le y_2 \le y_1 \le y_3$	$\frac{1}{2}[1-\sin\xi\cos\xi]$	$\frac{1}{2}[\sin^2\xi-\cos^2\xi]$
8	$y_4 \le y_1 \le y_3 \le y_2$	$\sin^2\xi + \frac{1}{2}\cos^2\xi$	$\cos\xi\sin\xi$

At the special points this gives the following derivatives

	left derivative	rightderivative	S* value	classification
0	0	0	1/2	saddle point
$1/4\pi$	-1/2	0	1/4	local minimum (global)
$1/2\pi$	1/2	-1/2	1/2	local maximum
$3/4\pi$	0	$\sqrt{2}$	1/4	local minimum (global)
π	0	0	1/2	saddle point
$5/4\pi$	$\sqrt{2}$	0	3/4	local maximum (global)
$8/2\pi$	-1/2	1/2	1/2	local minimum
$7/4\pi$	0	-1/2	3/4	local maximum (global)
2π	0	0	1/2	saddle point

Some additional information is contained in the enclosed memo. It is obvious from this example that the derivatives of S* may not exist at the minimum, even for linear models. I think this sufficiently explains the zigzagging behavior of the SSA programs when they approach convergence. The example also shows how the purpose of SSA can not be formulated as solving the stationary equation $\partial S^* = 0$, but it must be formulated as the minimization of S*. In general these two problems are quite different. In the example all stationary points are actually saddle points. This means that we must use the more general theory for the minimization of directionally differentiable functions (also called nonsmooth functions) which has been developed in recent years, mainly by Soviet mathematicians, and both from a theoretical and practical point of view. I hope these remarks are of some use to you, and I hope you don't mind my sending a copy of this letter to some people who may be interested.

Best regards,

Jan

J. DeLeeuw

MH-1229-JD-mm

Copy to
See next page

Copy to:
Dr. J. D. Carroll
Dr. J. B. Kruskal
Prof. J. Lingoes
Prof. E. I. Roskam
Prof. R. N. Shepard
Prof. F. W. Young

THE UNIVERSITY OF MICHIGAN
COMPUTING CENTER
1075 BEAL AVENUE
ANN ARBOR, MICHIGAN 48105

April 10, 1974

Dr. Jan DeLeeuw
Bell Telephone Laboratories
600 Mountain Ave.
Murray Hill, New Jersey 07974

Dear Jan:

Reur letter to Louis anent "Smoothness properties...": my first general comment is that special examples such as you have illustrated while of theoretical interest, do not have much relevance to the multivariable case addressed by our various algorithms. A second observation relates to your formulation of the problem for rank images as being essentially a single-phase algorithm à la Kruskal- neglecting the strong importance of the double-phase G-L algorithm, for which your comments are irrelevant. Neither I, nor Louis, nor Eddy have much (if any) faith in the d*-single-phase solution!

I'm not sure what you mean by "zigzagging", but that is not the way I would have characterized the behaviour of the G-L algorithm.

Relating the above comments to a misconception advanced at the bottom of p. 5, i.e., ...minimizing S* for fixed Øi(Θ)", I would refer you to the monograph (shortly to appear in Psychometrika) by myself and Roskam, which tries to clarify the various algorithmic distinctions.

I would hate to believe that after this heroic attempt on our part that "we all" would once more be engaged in a "correspondence musical chairs" on these issues (although I must admit that my comments were unsolicited).

Thank you again for keeping me abreast of your work, which I have followed for some time with great interest. If you have a copy of your thesis I would be most grateful to obtain it.

Sincerely,

James C. Lingoes
James C. Lingoes
Assoc. Prof., Dept. of Psychology
& Assoc. Res. Scientist, Computing Center

Copy to: Dr. J.D. Carrol
Dr. J.B. Kruskal
Prof. E.I. Roskam
Prof. R.N. Shepard
Prof. F.W. Young

JCL/dc