Bounds for SFORM1

Part I

Jan de Leeuw

RB 001 - 73

Summary

Monte Carlo studies using Kruskal's Stress (formula one) have shown that the maximin value of S (maximum over all possible data structures, minimum over all possible configurations) is strictly less than unity. In this paper we give a rigorous proof of the fact that maximin $S < \frac{1}{3}\sqrt{3}$, and the existing Monte Carlo results show that this bound certainly cannot be improved much, and is probably sharp. (even if we take the maximum over the much smaller class of <u>untied</u> data structures). Improvements of the bound which take into account the number of points and the number of dimensions n and p are considered. From the results of this paper we conjecture that maximin $S_n^p = \beta(p) - \delta(n,p)$, with $\delta(n,p)$ an increasing, nonnegative function of order $O(n^{-1})$ for each fixed p, and $\beta(p)$ a decreasing function of p of order $O(p^{-\frac{1}{2}})$.

m = 1 opt.

$$\sqrt{1 - \frac{1}{2}\left(\frac{n-1}{n}\right)^2}$$

Kruskal's Stress (formula one) is given by

$$S_n^p(X) = \min_{\hat{d}} \sqrt{\frac{\sum\sum(d_{ij} - \hat{d}_{ij})^2}{\sum\sum d_{ij}^2}},$$

where the d_{ij} are distances between the endpoints of n p-dimensional vectors whose coordinates are collected in the n x p matrix X, and where the $\hat{d}_{ij}$ must satisfy linear inequality restrictions of the form

$$\delta_{ijkl}(\hat{d}_{ij} - \hat{d}_{kl}) \geqslant 0.$$

Summation in the formula for S is over all $1 \leqslant i < j \leqslant n$, the signature δ_{ijkl} in the inequality constraints is a given set of real numbers. Clearly the vector $\tilde{d}_{ij}$, with all elements equal to the average $\bar{d}$ of the $\binom{n}{2}$ distances, satisfies these order restrictions. Consequently

$$S_n^p(X) \leqslant T_n^p(X) \triangleq \sqrt{\frac{\Sigma\Sigma\,(d_{ij} - \bar{d})^2}{\Sigma\Sigma d_{ij}^2}}, =$$ $$\sqrt{\Sigma d^2 - N\,\bar{d}\,\frac{(\Sigma d)^2}{N\,\Sigma d^2}}$$

and

$$\min_X S_n^p(X) \leqslant \min_X T_n^p(X),$$

where the minimum is taken over all n x p configuration matrices X. This is the basic inequality we start with. It implies, that, for any configuration matrix Y,

$$\min_X S_n^p(X) \leqslant T_n^p(Y).$$

In this paper we compute $T_n^p(Y)$ for some interesting special cases, and we make some general comments on the problem of minimizing $T_n^p(X)$. Of course T is the ratio of the standard deviation and the root mean square of the distances. It is a coefficient of variation, and configurations with small values of T must have their distances as equal as possible (in fact for the regular simplex in n - 1 dimensions the value of T_n^{n-1} is equal to zero). Consequently for our special cases we investigate regular configurations with large tie-blocks of distances.

For n equally spaced points in one dimension we find

$$\sum_{j=1}^{n}\sum_{i=1}^{j} d_{ij}^2 = \tfrac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n} d_{ij}^2 = \tfrac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{n}(i-j)^2 = n\sum_{i=1}^{n} i^2 - \left(\sum_{i=1}^{n} i\right)^2 =$$

$$= \frac{1}{6}n^2(n+1)(2n+1) - \frac{1}{4}n^2(n+1)^2 = \frac{1}{12}n^2(n+1)(n-1),$$

and

$$\sum_{j=1}^{n}\sum_{i=1}^{j} d_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{j}(j-i) = \sum_{j=1}^{n}\left[j^2 - \tfrac{1}{2}j(j+1)\right] =$$

$$= \tfrac{1}{2}\left[\frac{1}{6}n(n+1)(2n+1) - \tfrac{1}{2}n(n+1)\right] = \frac{1}{6}n(n+1)(n-1).$$

Thus

$$\bar{d} = \frac{1}{3}(n+1),$$

and substitution of these results in the formula for T_n^p gives

$$T_n^1(Y) = \sqrt{\frac{n-2}{3n}} = \frac{1}{3}\sqrt{3}\,\left(1 - \frac{2}{n}\right)^{\frac{1}{2}} = \frac{1}{3}\sqrt{3} + O(n^{-1}).$$

Consequently $T_n^1(Y)$ increases with n to the limit $\frac{1}{3}\sqrt{3}$. We have proved the chain

$$\min_X S_n^p(X) \leqslant \min_X S_n^1(X) \leqslant \min_X T_n^1(X) \leqslant T_n^1(Y) = \sqrt{\frac{n-2}{3n}} < \frac{1}{3}\sqrt{3}.$$

An important implication is that if we apply Kruskal's methodology to a complete set of dissimilarities (using stress formula one), then any stationary value larger then $\frac{1}{3}\sqrt{3} \approx .57735$ certainly corresponds with a local minimum. This is true for $p = 1$, and it is a fortiori true for $p > 1$. It is also irrelevant which one of the power metrics we use to compute d_{ij}.

The computations are somewhat more complicated in the case of n points equally spaced on a circle (using Euclidean distances). In the first place

$$\sum_{j=1}^{n} \sum_{i=1}^{j} d_{ij}^2 = 2 \sum_{j=1}^{n} \sum_{i=1}^{j} (1 - \cos \frac{j-i}{n} 2\pi) = \sum_{i=1}^{n} \sum_{j=1}^{n} (1 - \cos \frac{j-i}{n} 2\pi) =$$

$$= n^2 - (\sum_{i=1}^{n} \cos \frac{2\pi}{n} i)^2 (\sum_{i=1}^{n} \sin \frac{2\pi}{n} i)^2.$$

By taking forward differences and using familiar goniometrical identities

$$\sin \frac{2\pi}{n} (i+1) - \sin \frac{2\pi}{n} i = 2 \cos \frac{2\pi}{n} \frac{(i+1)+i}{2} \sin \frac{2\pi}{n} \frac{(i+1)-i}{2} =$$

$$= 2 \sin \frac{\pi}{n} (\cos \frac{2\pi}{n} i \cos \frac{\pi}{n} - \sin \frac{2\pi}{n} i \sin \frac{\pi}{n}),$$

$$\cos \frac{2\pi}{n} (i+1) - \cos \frac{2\pi}{n} i = -2 \sin \frac{2\pi}{n} \frac{(i+1)+i}{2} \sin \frac{2\pi}{n} \frac{(i+1)-i}{2} =$$

$$= -2 \sin \frac{\pi}{n} (\sin \frac{2\pi}{n} i \cos \frac{\pi}{n} + \cos \frac{2\pi}{n} i \sin \frac{\pi}{n}).$$

By summation of these differences we find the relations

$$\cos \frac{\pi}{n} \sum_{i=0}^{n-1} \cos \frac{2\pi}{n} i - \sin \frac{\pi}{n} \sum_{i=0}^{n-1} \sin \frac{2\pi}{n} i = 0,$$

$$\sin \frac{\pi}{n} \sum_{i=0}^{n-1} \cos \frac{2\pi}{n} i + \cos \frac{\pi}{n} \sum_{i=0}^{n-1} \sin \frac{2\pi}{n} i = 0.$$

The determinant of this linear system is equal to

$$(\cos \frac{\pi}{n})^2 + (\sin \frac{\pi}{n})^2 = 1,$$

and consequently

$$\sum_{i=0}^{n-1} \sin \frac{2\pi}{n} i = \sum_{i=0}^{n-1} \cos \frac{2\pi}{n} i = 0.$$

This implies

$$\sum_{i=1}^{n} \sin \frac{2\pi}{n} i = \sum_{i=1}^{n} \cos \frac{2\pi}{n} i = 0,$$

and thus

$$\sum_{j=1}^{n} \sum_{i=1}^{j} d_{ij}^2 = n^2.$$

Moreover

$$\sum_{j=1}^{n} \sum_{i=1}^{j} d_{ij} = \sqrt{2} \sum_{j=1}^{n} \sum_{i=1}^{j} \sqrt{1 - \cos \frac{j-i}{n} 2\pi} = 2 \sum_{j=1}^{n} \sum_{i=1}^{j} \sin \frac{j-i}{n} \pi =$$

$$= 2 \sum_{k=1}^{n-1} k \sin \frac{n-k}{n} \pi = 2 \sum_{k=0}^{n-1} k \sin \frac{k}{n} \pi.$$

Using the same methods as before we find

$$\cos\frac{\pi}{2n}\sum_{k=0}^{n-1}\cos\frac{\pi}{n}k - \sin\frac{\pi}{2n}\sum_{k=0}^{n-1}\sin\frac{\pi}{n}k = 0,$$

$$\sin\frac{\pi}{2n}\sum_{k=0}^{n-1}\cos\frac{\pi}{n}k + \cos\frac{\pi}{2n}\sum_{k=0}^{n-1}\sin\frac{\pi}{n}k = 1/\sin\frac{\pi}{2n}.$$

The solution of this system is

$$\sum_{k=0}^{n-1}\cos\frac{\pi}{n}k = 1,$$

$$\sum_{k=0}^{n-1}\sin\frac{\pi}{n}k = \operatorname{ctg}\frac{\pi}{2n}.$$

Summation by parts gives the formulae

$$\sum_{k=0}^{n-1}k\left(\sin\frac{k+1}{n}\pi - \sin\frac{k}{n}\pi\right) = -\operatorname{ctg}\frac{\pi}{2n},$$

$$\sum_{k=0}^{n-1}k\left(\cos\frac{k+1}{n}\pi - \cos\frac{k}{n}\pi\right) = -(n-1),$$

and our previous methods give the linear system

$$2\sin\frac{\pi}{2n}\cos\frac{\pi}{2n}\sum_{k=0}^{n-1}k\cos\frac{\pi}{n}k - 2\sin\frac{\pi}{2n}\sin\frac{\pi}{2n}\sum_{k=0}^{n-1}k\sin\frac{\pi}{n}k = -\operatorname{ctg}\frac{\pi}{2n},$$

$$2\sin\frac{\pi}{2n}\sin\frac{\pi}{2n}\sum_{k=0}^{n-1}k\cos\frac{\pi}{n}k + 2\sin\frac{\pi}{2n}\cos\frac{\pi}{2n}\sum_{k=0}^{n-1}k\sin\frac{\pi}{n}k = n-1.$$

The solutions are

$$\sum_{k=0}^{n-1}k\cos\frac{\pi}{n}k = \frac{1}{2}\left[\left(\operatorname{ctg}\frac{\pi}{2n}\right)^2 + (n-1)\right],$$

$$\sum_{k=0}^{n-1}k\sin\frac{\pi}{n}k = \frac{1}{2}n\operatorname{ctg}\frac{\pi}{2n},$$

$vp = n.$

and thus

$$\sum_{j=1}^{n}\sum_{i=1}^{j}d_{ij} = n\operatorname{ctg}\frac{\pi}{2n},$$

$$T_n^2(Y) = \sqrt{1 - \frac{2\left(\operatorname{ctg}\frac{\pi}{2n}\right)^2}{n(n-1)}}.$$

$$\boxed{\sqrt{1 - \frac{2\left[v^2 p \operatorname{ctg}\frac{\pi}{2p}\right]^2}{n(n-1)\; n^2}} = \sqrt{1 - \frac{2v^{[illegible]}\operatorname{ctg}\frac{\pi}{2p}}{p\left(p-\frac{1}{v}\right)}}}$$

Because

$$\operatorname{ctg}\frac{x}{n} = \frac{n}{x} + O(n^{-2})$$

we find

$$T_n^2(Y) = \sqrt{\frac{\pi^2 - 8}{\pi^2}} + O(n^{-1}).$$

The main difference with the previous example is that both the average distance $\bar{d}$ and the average squared distance $\bar{e}$ are bounded sequences. In fact

$$\bar{d} = \frac{4}{\pi} + O(n^{-1}),$$

$$\bar{e} = 2 + O(n^{-1}).$$

A table of $T_n^2(Y)$ is given below.

n	$T_n^2(Y)$	
2	0	
3	0	
4	0.169102	
5	0.229753	
6	0.267307	
7	0.293122	
8	0.312013	.30834
9	0.326452	.3209

n	$T_n^2(Y)$	
10	0.337854	.331482
15	0.371315	?.362
20	0.387654	
25	0.397339	?.386
30	0.403747	
40	0.411705	
50	0.416452	?.403
75	0.422749	

n	T_n^2
1E2	0.425885
1E3	0.434305
1E4	0.435143
1E5	0.435227
1E6	0.435235
1E7	0.435236

$\sqrt{1-\frac{2}{3}\frac{n}{n-1}}$ $\qquad\qquad$ $\sqrt{1-\frac{2}{3}\frac{n^2-u}{n^2-n}}$

		u	
5	.4082483	2	.5322907
10	.5091751	1	.5163970
15	.5345224	0	→
20	.5461187	2	.5493217
25	.5527708	1	.5537749
30	.5570860	0	→
40	.5623516	1	.5627314
∞	.5773502	∞	→

For $p > 2$ we do not have a complete sequence of geometrical examples, but we have computed $T_n^p(Y)$ for some of the regular polytopes discussed, for example, in Sommerville (1929/1958, p 179-185). The results are

$T_5^3 = .127813$

$p = 3$	
tetrahedron	$T_4^3 = 0$
octahedron	$T_6^3 = 0.151249$
cube	$T_8^3 = 0.202805$
icosahedron	$T_{12}^3 = 0.242322$
dodecahedron	$T_{20}^3 = 0.279158$

$p = 4$	
5-cell	$T_5^4 = 0$
16-cell	$T_8^4 = 0.135583$
8-cell	$T_{16}^4 = 0.211457$
24-cell	$T_{24}^4 = 0.229238$
600-cell	$T_{120}^4 = 0.$

Using a rather conservative extrapolation rule we conjecture that

$$\min_X T_n^3(X) \leqslant .333,$$

$$\min_X T_n^4(X) \leqslant .265,$$

for all n. This can be compared with our previous results

$$\min_X T_n^2(X) \leqslant .436,$$

$$\min_X T_n^1(X) \leqslant .578.$$

$$C_p = \frac{2^{p-2}\,\Gamma^2(\frac{1}{2}p)}{\Gamma(\frac{1}{2})\,\Gamma(p-\frac{1}{2})} = \frac{2^{p-2}\,\Gamma^2(\frac{1}{2}p)}{\sqrt{\pi}\,\Gamma(p-\frac{1}{2})}$$

$$2C_p^2 = \frac{2^{2p-4}\cdot\Gamma^4(\frac{1}{2}p)}{\pi\,\Gamma^2(p-\frac{1}{2})} \quad ?$$

In the previous sections we have studied examples in which p was fixed and in which n varied. We now consider an example in which n and p vary together. The p-dimensional octahedron has $n = 2p$ vertices with coordinates $\pm e_i$, where e_i are the p-dimensional unit vectors (the rows of the identity matrix of order p). This obviously implies that among the d^2_{ij} with $1 \leqslant i < j \leqslant n$ there are $\frac{1}{2}n$ elements equal to four and $\frac{1}{2}n(n-1) - \frac{1}{2}n = \frac{1}{2}n(n-2)$ elements equal to two.

Thus

$$\sum_{j=1}^{n} \sum_{i=1}^{j} d_{ij} = \tfrac{1}{2}\sqrt{2}\, n(n-2) + n,$$

$$\sum_{j=1}^{n} \sum_{i=1}^{j} d^2_{ij} = n(n-2) + 2n = n^2,$$

$$\sum\sum d_{ij}^4 = 2n(n-2) + 8n = 2n(n+2)$$

$$T_n^{\frac{1}{2}n}(Y) = \sqrt{3 - 2\sqrt{2}} \cdot \sqrt{\frac{1}{n}} \cdot \sqrt{\frac{n-2}{n-1}} = O(n^{-\frac{1}{2}}).$$

Consequently T converges (to zero), but convergence is slower than in the examples with p fixed. Observe

$\overline{d} = \sqrt{2} + O(n^{-1})$.

Convergence of T to zero is due to the fact that almost all distances are equal to $\sqrt{2}$ (the proportion of distances not equal to $\sqrt{2}$ is $O(n^{-1})$ as well).

SStress $\sqrt{1 - \dfrac{\cancel{2}\, n^{2}}{n(n-1)\, \cancel{2}n(n+2)}}$

$$= \sqrt{1 - \frac{n^2}{(n-1)(n+2)}}$$

$$\sqrt{1 - \left(\frac{n}{n-1}\right)\left(\frac{p}{p+1}\right)}$$

$$= \sqrt{1 - \left(\frac{n}{n-1}\right)\frac{\frac{1}{2}n}{\frac{1}{2}n+1}}$$

$$\sqrt{1 - \left(\frac{n}{n-1}\right)\left(\frac{n}{n+2}\right)}$$

In the previous section we studied an example in which both n and p varied in such a way that their rates of increase were equal. We now study a case in which n increases faster than p (if fact faster than any power of p). Consider the $n = 2^p$ vertices of the unit hypercube in p dimensions. The n^2 squared distances assume the values 0, 1, ... , p with frequencies $\binom{p}{0}2^p$, $\binom{p}{1}2^p$, ..., $\binom{p}{p}2^p$. Thus

$$\sum_{j=1}^{n}\sum_{i=1}^{j} d_{ij}^2 = 2^{p-1}\sum_{k=0}^{p}\binom{p}{k} k.$$

$K(2^p, p) \leq$

From the theory of the binomial distribution

$$\frac{1}{2} = \sum_{k=0}^{p}\binom{p}{k}\left(\tfrac{1}{2}\right)^p\left(\frac{k}{p}\right) = \frac{1}{p2^p}\sum_{k=0}^{p}\binom{p}{k} k,$$

and thus

$$\sum_{j=1}^{n}\sum_{i=1}^{j} d_{ij}^2 = p\ 2^{2p-2}.$$

For the sum of the distances we find

$$\sum_{i=1}^{n}\sum_{j=1}^{j} d_{ij} = 2^{p-1}\sum_{k=0}^{p}\sqrt{k}\binom{p}{k}.$$

A closed form for this sum does not seem to exist. For the asymptotics we define

$$A(p) = \sum_{k=0}^{p}\sqrt{\frac{k}{p}}\binom{p}{k}\left(\tfrac{1}{2}\right)^p,$$

$$B(p) = \frac{2^{n+1}}{2^n - 1}.$$

By substitution we find

$$T_n^p(Y) = \sqrt{1 - A^2(p)B(p)}.$$

Of course B(p) decreases to 2 faster than any power of p. To evaluate the limit of A(p) we interpret it as the expected value of the square root of a proportion in a Binomial model with probability of success equal to $\frac{1}{2}$. From standard large sample theory

$$A(p) = \tfrac{1}{2}\sqrt{2} + O(p^{-1}),$$

Combining these results

$$T_n^p = O(p^{-\frac{1}{2}}).$$

The following table seems useful.

p	A(p)	B(p)	$T_n^p(Y)$
1	0.5	4	0
2	0.603553	2.666667	0.169102
3	0.647692	2.285714	0.202805
4	0.669171	2.133333	0.211457
5	0.680585	2.064515	0.209102
6	0.687134	2.031746	0.201757
7	0.691173	2.015748	0.192452
8	0.693839	2.007843	0.182758
9	0.695709	2.003914	0.173449
10	0.697091	2.001955	0.164862
20	0.702155	2.000002	0.114509
30	0.704062	2.000000	0.092706
40	0.704843	2.000000	0.079963
50	0.705304	2.000000	0.071353
100	0.706214	2.000000	0.050223

Observe that we have not developed the asymptotics in terms of the number of points, but in terms of the number of dimensions. The impression that the convergence in this example is of the same order as the convergence in the previous example is completely wrong. In fact we find that scaling 2^{100} (i.e. more that 10^{30}) points in 100 dimensions can still give stress values of 0.05.

-5-

Up to now we have established a number of bounds based on the inequalities

$$\min_X S_n^p(X) \leq \min_X T_n^p(X) \leq T_n^p(Y).$$

We did not attempt any systematic theory for fixed $p > 2$, mainly because we did not have a satisfactory theory of equal spacing in $p > 2$ dimensions (I would conjecture, however, that a more systematic use of nonlinear coordinate systems such as multidimensional polar coordinates would make such attempts possible. Moreover for none of our choices of Y have we proved that

$$\min_X T_n^p(X) = T_n^p(Y),$$

and we have not at all considered the question whether

$$\max \min_X S_n^p(X) = \min_X T_n^p(X),$$

where the maximum is taken over all possible signatures σ_{ijkl}. In this section we give some fragmentary results on these questions.

If the class of signatures we admit consists of all possible four dimensional structures of real numbers, then any signature with $\sigma_{ijkl} = \sigma_{klij} > 0$ for all i,j,k,l implies that $S_n^p(X) = T_n^p(X)$ for all X. For this class of signatures bound based on T_n^p are sharp (in multidimensional scaling terminology this is the case in which the data contain arbitrary many ties and we use semi-strong monotonicity, or, equivalently, the secondary approach). If $\sigma_{ijkl} = -\sigma_{klij}$ and $\sigma_{ijkl} \neq 0$ for all i,j,k,l (if the data do not contain ties), then the computations of Lingoes and Roskam (1971, table 10) show that

$$\min_X S_4^1(X) < 0.170,$$

while the analysis in section 2 gives

$$\min_X T_4^1(X) < 0.409,$$

which is quite a long way off. It seems obvious however that such a large discrepancy is at least partly due to the small number of points, and other evidence (from MonteCarlo studies of Spence, Wagenaar and Padmos, and others) indicates that if n increases the bounds may very well become sharp.

More precisely, we can derive the following table for max min $S_n^p(X)$ from the work of Mahr (1969, table 1). Entry (n,p) in the table is actually a lower bound for this value.

n \ p	1	2	3	4
6	0.471	0.218	0.096	0.000
7	0.477	0.212	0.101	0.025
8	0.470	0.260	0.122	0.079
10	0.498	0.267	0.147	0.096
12	0.445	0.288	0.178	0.131
16	0.505	0.300	0.207	0.152
∞	0.56	0.38	0.30	0.24

The lower bound in the final row is derived from Stenson & Knoll (1969). Comparison with the values we computed in the previous section shows that our bounds cannot be improved much, especially for larger values of n (or large values of n/p). Asymptotically the bounds seem to be sharp.

Another useful way to study the behaviour of the function $T_n^p(X)$ is to compute derivatives and study the first order conditions for a stationary value. Of course minimization of T_n^p is equivalent to maximization of

$$[illegible] = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij},$$

under the condition that

$$\sum_{i=1}^{n} \sum_{s=1}^{p} x_{is}^2 = 1,$$

$$\sum_{i=1}^{n} x_{is} = 0 \text{ for all } s = 1, \ldots, p.$$

Using only the first (quadratic) restrictions, we find the conditions for a stationary value. They can be written as the matrix equation

$$AX = \lambda X,$$

where $A = \{a_{ij}\}$ is defined by

$$a_{ij} = \begin{cases} -d_{ij}^{-1} & \text{if } i \neq j, \\ \sum_{j \neq i}^{n} d_{ij}^{-1} & \text{if } i = j. \end{cases}$$

It follows that A is positive semidefinite of rank $\leqslant n - 1$ for all possible choices of X. Moreover A is doubly centered, which means that we can suppose without loss of generality that X satisfies the second set of conditions we did not use in the derivation. The optimal A must have an eigenvalue of multiplicit p, and consequently it is not very surprising that the optimal A gives regular patterns of distances.

Although we could use these stationary equations to compute optimal values of T_n^p, and to check if the configurations we have chosen correspond with stationary vales, wo do not do this systematically. It is comparatively easy to check that most of our configurations do satisfy the conditions, but this does not prove that they correspond with absolute minima. Perhaps more satisfactory results could be obtained if we use what little convexity there is in the problem to obtain a more satisfactory characterization of the stationary values, perhaps useful results could also be derived from the second order conditions. Again, as in section 7, we remark that progress could possible be made by abandoning the Cartesian coordinates, and by studying the problem on nonlinear manifolds.

It is, however, interesting to observe that in the one-dimensional case the situation is basically much more simple. We suppose, as in section 2, that $x_1 < x_2 < \dots < x_n$. Now

$$a_{ij} = \begin{cases} (x_i - x_j)^{-1} & \text{if } i > j, \\ (x_j - x_i)^{-1} & \text{if } i < j. \end{cases}$$

Consequently

$$\sum_{j=1}^{n} a_{ij}x_j = -\sum_{j=1}^{i-1} \frac{x_j}{x_i - x_j} + \sum_{j=1}^{i-1} \frac{x_i}{x_i - x_j} + \sum_{j=i+1}^{n} \frac{x_i}{x_j - x_i} - \sum_{j=i+1}^{n} \frac{x_j}{x_j - x_i} =$$

$$= (i - 1) - (n - i) = 2i - (n + 1).$$

Consequently the unique solution of the stationary equations in the case $p = 1$ is obtained by taking x proportional to the centered rank numbers, i.e. equally spaced. This corresponds with the absolute maximum of $H_n^1(X)$, and we have proved

$$\min_X T_n^1(X) = \sqrt{\frac{n-2}{3n}} \uparrow \frac{1}{3}\sqrt{3}.$$

ook zoals
voor SStress

At least two practical applications of these results are immediately obvious. It is useful to know the 'effective' range of a coefficient one is minimizing, the idea that

$$0 \leq \min_X S_n^p(X) \leq 1$$

can be quite misleading. In Monte Carlo research it has already been conjectured that max min $S_n^p(X) < 1$ (Wagenaar and Padmos, 1971, p 108), the results in this paper give a rigorous proof of this fact.

In the second place the idea that SFORM1 may be systematically biassed towards equally spaced configurations in the no-structure case may be interpreted as a disadvantage. In this context we observe that Stress (formula 2) can be defined as

$$\underline{S}^p(X) = \frac{S_n^p(X)}{T_n^p(X)} .$$

This shows that the alternative MDSCAL loss-function avoids all possible difficulties of this type by definition. In 1969, when the result for $p = 1$ was discovered (De Leeuw 1970), I concluded that SFORM2 should be preferred to SFORM1 because of this effect. This seems a somewhat too hasty conclusion now. As far as there is real structure in the data the distorting effect will be slight.

Other Monte Carlo results can also be explained (or made more easily understandable) by our results. Spence (1970, p 69) dicovered, for example, that the mean of SFORM2 (for given level of error, and for given number of points) did not change as much as the mean of SFORM1 when the dimensionality of the space was increased.

Klahr, D.

A Monte Carlo investigation of the statistical significance of Kruskal's nonmetric scaling procedure.

Psychometrika, 34, 1969, 319-330.

Stenson, H.H. and Knoll, R.L.

Goodness of fit of random rankings in Kruskal' nonmetric scaling procedure.

Psychological Bulletin, 71, 1969, 122-126.

Lingoes, J.C. and Roskam, E.

A mathematical and empirical study of two multidimensional scaling algorithms.

Michigan mathematical psychology program, report MMPP 71-1, University of Michigan, 1971

Sommerville, D.M.Y.

An introduction to the geometry of n dimensions

1921: repreinted edition, New York, Dover, 195

De Leeuw, J.

The positive orthant method for nonmetric multidimensional scaling.

Department of Data Theory for the Social Scienc report RN 001-70, University of Leiden, 1970.

Spence, I.

Multidimensional scaling: an empirical and theoretical investigation.

Unpublished Ph.D. thesis, University of Toronto, 1970.

Wagenaar, W.A. and Padmos, P.

Quantitative interpretation of stress in Kruskal's multidimensional scaling technique.

Br. J. math. statist. Psychol., 24, 1971, 101-1

	CUBE $\binom{n=8}{p=3}$	octahedron $\binom{n=6}{p=3}$	Dodecahedron (
Σd	71.79753196	22.97056275	370.9429664
Σd^2	192	36	785.410[illegible]966
Σd^4	1536	96	4112.46118
$\Sigma \ln d$	25.76422872	6.238324625	117.7177236
$\Sigma (\ln d)^2$	24.9124286	2.882718084	93.41506219
$\binom{n}{2}$	28	15	190
σ_K	= .2028051073	.1512494078	.2791584907
σ_T	= .377964473	.316227766	.4508314677
σ_R	= 1.205447117	.2882718084	20.48104927

	Isocahedron $\binom{n=12}{p=3}$	8-cell $\binom{n=16}{p=4}$	16 cell $\binom{n=8}{p=4}$
Σd	89.95369786	342.6157537	41.9411255
Σd^2	130.2492236	1024	64
Σd^4	314.1640786	10240	160
$\Sigma \ln d$	18.2941471	122.9361681	11.09035489
$\Sigma (\ln d)^2$	9.427371579	132.0360069	4.804530139
$\binom{n}{2}$	66	120	28
σ_K	.2423215063	.21145664	.1355832145
σ_T	.4264014327	.3829708431	.2927700219
σ_R	4.356525852	6.091828374	.4118168691