RN 002-70

THE EUCLIDEAN DISTANCE MODEL

Jan de Leeuw

january 1970 - august 1970

DEPARTMENT OF DATA THEORY - UNIVERSITY OF LEYDEN - NETHERLANDS

Introduction & summary

This paper is a rough draft. This does not mean that I don't want people to quote it, or even that I don't want people to read it. It means that right now I should like to organize it differently if it would not take so much time, and that there are some sections which I consider weak and quite useless.

There is enough material in this paper for five or more articles of the usual journal size. But I don't think any of it is going to be published in that form (at least not by me). The non-redundant parts will be included in my forthcoming dissertation.

In my point of view the most important parts of the paper are:

1) The NMSEMS or CARD9A method discussed in chapter 1. From a practical point of view this is by far the most important contribution, especially when we consider the straightforward extensions of this method to all four kinds of Coombsian data matrices, to three-way data structures, and to categorical (binary) data.
2) The investigation of the structure of the solution set, its mathematical properties.
3) The discussion of uniqueness.
4) The generalization and the QP-formulation of the AT-method
5) The computational suggestions for the AT- and LS-methods.
6) The representation theorems in appendix B.
7) The terminology used throughout this paper (at some places the notation is sloppy and will be improved).

Again from a practical point of view the main conclusions are as follows. There are four different types of algorithms for the NMDS-problem. The first one is NMSEMS. It has the disadvantage that its loss-function is not too satisfactory, and that it may not find a perfect solution, even in n-1 dimensions. The advantages are: no local minima, if it finds a solution in the cone it will find a representative one. The second class contains the projection-type algorithms of which MINISSA is the best example. Advantages: relatively fast, parsimonious solution, nice loss-function. Disadvantages: the most serious one is degeneracy, often disguised as weak order-isometricity. The third one is the LS-solution, which can only be applied with numerical dissimilarities, and the final one is the AT-solut which takes an awful lot of time and generally gives a solution in not less than n - 1 dimensions. In some cases it may be the only reasonable possibili and we may always round the solution to $p < n - 1$ dimensions using a princip component error theory.

There are only a few numerical examples in this paper, and their size is very unrealistic. For NMSEMS quite a number of programs are ready, and I am busy writing a program for AT, because that interests me most. Quite a job, howev

I should appreciate very much receiving comments and criticisms of this paper. My present address is

Jan de Leeuw

Department of data theory for the social sciences

c.a. Centraal Reken Instituut ; room 101a

Stationsplein 10

Leiden

Netherlands

Note: The symbol '$\xrightarrow{D\ k}$' in the margin indicates that note k in appendix D is relevant at this place.

Contents

0 Preliminaries 1

1 A quick-and-dirty method 3

2 Existence and uniqueness

2.1 The set of solutions 8
2.2 Convexifying the problem 9
2.3 Representation theorem 11
2.4 Uniqueness question 12

3 Linearizing the problem

3.1 Sufficient conditions 16
3.2 The T-method 17
3.3 Necessary conditions 21
3.4 The complete method 21
3.5 Computational remarks 25
3.6 Initial spanning set 27
3.7 A somewhat different approach 35
3.8 The role of nI - E 37
3.9 Further computational suggestions 47

4 Selecting a particular solution

4.1 Optimal p-parsimoneous solutions 55
4.2 Minimum deviation solutions 57
4.3 Degrees of faithfulness 61
4.4 Representative solutions 66

5 A comparison of different algorithms 82

Appendices

A Additional definitions 89
B A summary of representation theorems 93
C Advanced computational remarks 98
D Corrections, Addenda, Remarks 106

References 111

0 Preliminaries

The primitives in our problem are an n-element stimulus-set A, a single valued mapping ϕ of A x A into the set of dissimilarities $\Delta \equiv \phi(A \times A)$, and a partial order $\geqslant_o$ over the elements of Δ. The elements of A are written as a_i $(i=1,\dots,n)$, the elements of Δ as $\delta_{ij} = \phi(a_i, a_j)$. The index set consisting of the first n integers is denoted by N, the asymmetric and symmetric subsets of $\geqslant_o$ by $>_o$ and $=_o$.

D:

Definition 0.1: The function ϕ is called a partially ordered generalized (or POG) distance iff the following conditions are satisfied for all $i,j,k,l \in N$

$$0.1.1:\quad \delta_{ij} \geqslant_o \delta_{ji} \quad \text{iff} \quad \delta_{ji} \geqslant_o \delta_{ij},$$

$$0.1.2:\quad \text{not } \delta_{ii} >_o \delta_{jk},$$

$$0.1.3:\quad \text{if } \delta_{ii} \geqslant_o \delta_{jk} \text{ then for all } l \ \delta_{jl} =_o \delta_{kl}.$$

Let χ be a mapping of A into R^p (the set of all p-tuples of real numbers). The set $\chi(A)$ is called a (p-dimensional) representation of A. The elements of $\chi(A)$ are p-element vectors called points and denoted by x_i, the coordinate values are the real numbers x_{is} $(i=1,\dots,n\ ;\ s=1,\dots,p)$. The coordinate values can be collected in an n x p matrix $X = \{x_{is}\}$, called the configuration matrix. The symbol d_{ij} is used for the Euclidean distance between x_i and x_j.

Definition 0.2: A representation $\chi(A)$ is strongly order-isometric if for all $i,j,k,l \in N$

$$0.2.1\quad \delta_{ij} >_o \delta_{kl} \Rightarrow d_{ij} > d_{kl},$$

$$0.2.2\quad \delta_{ij} =_o \delta_{kl} \Rightarrow d_{ij} = d_{kl}.$$

Such a representation in p dimensions is also called a strong p-representation.

Definition 0.3: A representation $\chi(A)$ is semi-strongly order-isometric if for all $i,j,k,l \in N$

$$0.3.1: \quad \delta_{ij} \geq_o \delta_{kl} \Rightarrow d_{ij} \geq d_{kl} ,$$

$$0.3.2: \quad \delta_{ij} =_o \delta_{kl} \Rightarrow d_{ij} = d_{kl} .$$

Such a representation in p dimensions is also called a <u>semi-strong p-representation</u>.

<u>Definition 0.4</u>: A representation $\chi(A)$ is <u>weakly order-isometric</u> if for all $i,j,k,l \in N$

$$0.4.1: \quad \delta_{ij} \geq_o \delta_{kl} \Rightarrow d_{ij} \geq d_{kl} .$$

Such a representation in p dimensions is also called a <u>weak p-representation</u>.

Clearly any strong p-representation is a semi-strong p-representation, and any semi-strong p-representation is a weak p-representation. The (p dimensional) Euclidean nonmetric multidimensional scaling (ENMS) problem is to find an order-isometric representation (in p dimensions), or, more generally, to find a (p-dimensional) representation which fits the requirements of order-isometricity optimally. The three different definitions of order-isometricity give us three different types of requirements.

1 <u>A quick-and-dirty method</u>

D6

D11

We define (using a term of Guttman) the <u>signature</u> of the data as the functi

$$\sigma_{ijkl} = \begin{cases} +1 & \text{if} \quad \delta_{ij} \geqslant_\delta \delta_{kl}, \\ -1 & \text{if} \quad \delta_{ij} \leqslant_\delta \delta_{kl}, \\ 0 & \text{otherwise.} \end{cases}$$

Then requiring weak order-isometricity is equivalent to requiring the existence of a solution to the set of n^4 homogeneous inequalities

$$\sigma_{ijkl}\,(d_{ij}^2 - d_{kl}^2) \geqslant 0. \tag{1}$$

The fact that we use the square of the Euclidean distance is not essential, of course, because $d_{ij} \geqslant 0$ for all i,j. Using the square slightly simplifie the computational problem in some cases, and, what is much more important, greatly simplifies the mathematical structure and analysis of our problem. Our quick-and-dirty method maximizes the sum of the left-hand terms of (1), subject to some scaling condition which makes the set of all permissible solutions bounded. Using the fact that $\geqslant_\delta$ is a partial order, it follows that $\sigma_{ijkl} = -\sigma_{klij}$, so

$$F = \sum_i \sum_j \sum_k \sum_l \sigma_{ijkl}(d_{ij}^2 - d_{kl}^2) =$$

$$2 \sum_i \sum_j d_{ij}^2 \sum_k \sum_l \sigma_{ijkl}.$$

Let

$$r_{ij} = \sum_k \sum_l \sigma_{ijkl},$$

then it is sufficient to maximize

$$G = \sum_i \sum_j d_{ij}^2 r_{ij}.$$

We introduce the set Σ_n of unit matrices of order n. The elements of this set, E^{ij}, are defined by

$$E^{ij} = \{e_{kl}^{ij}\} = \{\delta^{ki}\,\delta^{lj}\},$$

where superscripted δ is the Kronecker symbol. Thus E^{ij} has element e_{ij}^{ij} equal to unity, the other elements are zero. Evidently Σ_n has n^2 elements. Some related matrices we shall also need in the sequel are

$$E^{i\cdot} = \sum_j E^{ij} \quad ; \quad E^{\cdot j} = \sum_i E^{ij} \quad ; \quad E = \sum_i \sum_j E^{ij}.$$

Another class of matrices, denoted by $\mathcal{A}$, is defined by

$$A^{ij} = E^{ii} + E^{jj} - E^{ij} - E^{ji}.$$

The matrices A^{ii} are all equal to the null-matrix.

Theorem 1.1: If X is an $n \times p$ configuration matrix, then

$$d_{ij}^2 = \mathrm{Tr}(X'A^{ij}X).$$

Proof: The columns of X are $x_{.s}$ $(s=1,\dots,p)$. Then

$$\mathrm{Tr}(X'A^{ij}X) = \sum_s x'_{.s}A^{ij}x_{.s} = \sum_s (x_{is}^2 + x_{js}^2 - 2x_{is}x_{js}) =$$

$$\sum_s (x_{is} - x_{js})^2 = d_{ij}^2.$$

Q.E.D.

Using this theorem

D7 →

$$G = \mathrm{Tr}\left\{X'(\sum_i \sum_j r_{ij}A^{ij})X\right\} = \mathrm{Tr}(X'AX) \qquad \text{(say)}.$$

A natural scaling requirement is $X'X = I$, and the optimal solution is given by the eigenvectors corresponding with the p largest eigenvalues of A (evidently A is symmetric). We proceed to prove some properties of this solution.

Theorem 1.2: The column (and row) sums of A vanish. Consequently A is singular.

Proof: Let e denote an n-element vector with $e_i = 1$ for all $i \in N$. For all A^{ij} it is true that $A^{ij}e = 0$, consequently $\sum\sum r_{ij}A^{ij}e = 0$, or $Ae = 0$. Q.E.D.

Corrolary 1.1: A has an eigenvector with constant elements and eigenvalue zero. All other eigenvectors are centered.

Theorem 1.3: If $\geqslant$ is connected over Δ and ρ_{ij} denotes the rank number of δ_{ij} in the chain (ties are averaged, i.e. have equal rank numbers), then

$\frac{1}{2}r_{ij} = \rho_{ij} - \frac{1}{n^2}\sum\sum\rho_{ij}.$

Proof: $r_{ij} = \sum_k \sum_l \sigma_{ijkl}$. Thus r_{ij} is the number of elements of Δ that are strictly less than δ_{ij} (say n^-) minus the number of elements that are strictly greater than δ_{ij} (say n^+). Or: $r_{ij} = n^- - n^+$. The number of elements equal to δ_{ij} is $n^2 - n^+ - n^-$ ($= n^o$ say). The rank number of δ_{ij} is the average of the n^o numbers $n^-+1, \ldots, n^-+n^o$. This average equals

$$\rho_{ij} = \frac{n^o n^- + \frac{1}{2}n^o(n^o + 1)}{n^o} = n^- + \tfrac{1}{2}(n^o + 1) =$$

$$= n^- + \tfrac{1}{2}(n^2 - n^+ - n^- + 1) = \tfrac{1}{2}(n^- - n^+) + \tfrac{1}{2}(n^2 + 1).$$

Of course $\sum\sum\rho_{ij}$ is equal to the sum of the first n^2 integers, or

$$\frac{1}{n^2}\sum\sum\rho_{ij} = \tfrac{1}{2}(n^2 + 1),$$

and

$$\rho_{ij} - \bar{\rho} = \tfrac{1}{2}(n^- - n^+) = \tfrac{1}{2}r_{ij}.$$

Q.E.D.

<u>Theorem 1.4</u>: In general

$$a_{ij} = \delta^{ij}\sum_{k=1}^{n}(r_{ik} + r_{ki}) - (r_{ij} + r_{ji})$$

Proof: By definition $A = \sum\sum r_{ij}A^{ij}$. Consequently

$$a_{kk} = \sum_i \sum_j r_{ij}a^{ij}_{kk} = \sum_{i \neq j}\sum r_{ij}(\delta^{ik} + \delta^{jk}) =$$

$$= \sum_{j \neq k}(r_{kj} + r_{jk}) = \sum_j (r_{kj} + r_{jk}) - (r_{kk} + r_{kk}).$$

$$a_{kl} = \sum_i \sum_j r_{ij}a^{ij}_{kl} = -\sum_i \sum_j r_{ij}(\delta^{ik}\delta^{jl} + \delta^{jk}\delta^{il}) =$$

$$= -(r_{kl} + r_{lk}).$$

Combining these results gives the required formula. Q.E.D.

<u>Theorem 1.5</u>: If ϕ is a POG-distance and $\geq_o$ is connected (or: if ϕ is a weakly ordered generalized distance) then for all $i,j,k,l \in N$

1.5.1 $r_{ij} = r_{ji}$,

1.5.2 $r_{ii} = r_{jj}$,

1.5.3 $r_{ii} \leq r_{kl}$.

Proof: Definition 0.1.1 implies that $\delta_{ijkl} = \delta_{jikl}$ for all k,l, and thus $r_{ij} = \sum_k \sum_l \delta_{ijkl} = \sum_k \sum_l \delta_{jikl} = r_{ji}$. Applying the same reasoning with i = j proves that $r_{ii} = r_{jj}$. Definition 0.1.2 implies that $\delta_{iipq} \leq 0$ for all $p,q \in N$. If δ = 0, then δ = δ because $\geq$ is connected, and by definition 0.1.2 then $\delta_{kl} \geq \delta_{pq}$, or $\delta_{klpq} \geq \delta_{iipq} = 0$. If $\delta_{iipq} = -1$ then, by the definition of δ, $\delta_{klpq} \geq \delta_{iipq} = -1$. Summation over p,q proves $r_{kl} \geq r_{ii}$. Q.E.D.

If $\geq_c$ is connected we can construct the matrix B defined by

$$b_{ij} = \delta^{ij} \sum_{k=1}^{n} (\rho_{ik} + \rho_{ki}) - (\rho_{ij} + \rho_{ji}).$$

Theorem 1.6: B has the same eigenvectors as A.

Proof: According to theorem 1.3 we have $\frac{1}{2}r_{ij} = \rho_{ij} - \bar{\rho}$, where $\bar{\rho}$ is the (constant) average of the rank numbers. Applying theorem 1.4 gives

$$\begin{aligned}\tfrac{1}{2}a_{ij} &= \delta^{ij} \sum_{k} (\rho_{ik} - \bar{\rho} + \rho_{ki} - \bar{\rho}) - (\rho_{ij} - \bar{\rho} + \rho_{ji} - \bar{\rho}) = \\ &= b_{ij} - 2\delta^{ij} n \bar{\rho} + 2\bar{\rho} = \\ &= b_{ij} - 2\bar{\rho}\,(n\delta^{ij} - 1).\end{aligned}$$

Or

$$B = \tfrac{1}{2}A + 2\bar{\rho}\,(nI - E).$$

The row (and column) sums of both B and nI - E vanish, which means that e is an eigenvector with eigenvalue zero of both A and B. The other eigenvectors of A (and B) are centered, and for all centered vectors x it is true that $(nI - E)x = nx$. Thus, if $Ax = \lambda x$ with $x \neq e$, then

$$Bx = \tfrac{1}{2}Ax + 2\bar{\rho}(nI - E)x = (\tfrac{1}{2}\lambda + 2n\bar{\rho})x,$$

which means that x is also an eigenvector of B with eigenvalue $\frac{1}{2}\lambda + 2n\bar{\rho}$. Q.E.D.

Definition 1.1: A real symmetric matrix is called a T-matrix (after Taussky 1949) if

1.1.1 $\quad c_{ij} \leq 0 \qquad \forall i \neq j.$

1.1.2 $\quad \sum_k c_{ik} \geq 0 \qquad \forall i.$

Theorem 1.7: A T-matrix is

1.7.1 positive definite (PD) if $\sum_k c_{ik} > 0$ for all i,

1.7.2 positive semidefinite (PSD) if $\sum_k c_{ik} = 0$ for all i.

Proof: If C is a T-matrix then $\sum_j c_{ij} \geq 0$ or $c_{ii} + \sum_{j \neq i} c_{ij} \geq 0$ or $c_{ii} =$

$\Delta_i - \sum_{j \neq i} c_{ij}$ with $\Delta_i \geqslant 0$. Consider the inequality

$$0 \leqslant -\tfrac{1}{2} \sum_{i \neq j} \sum c_{ij}(x_i - x_j)^2 = \sum_i \sum_j c_{ij} x_i x_j - \sum_i x_i^2 \sum_j c_{ij} \leqslant$$

$$\leqslant \sum_i \sum_j c_{ij} x_i x_j .$$

If $\Delta_i > 0$ for all i and $x \neq 0$ then this inequality is strict. If $\Delta_i = 0$ for all i, then e is an eigenvector with eigenvalue zero. Q.E.D.

<u>Corrolary 1.2</u>: B is PSD. A is PSD iff A = 0 (this last part is true if all 'upper-diagonal' dissimilarities are weakly ordered).

Proof: If $i \neq j$ then $b_{ij} = b_{ji} \leqslant 0$, moreover $\sum_j b_{ij} = 0$ for all i. Thus B is a PSD T-matrix. If ϕ is a WOG-distance over the δ_{ij} with $j > i$, then $\sum_k a_{kk} = 0$, which means that the sum of the eigenvalues is zero. They are nonnegative iff they are all zero. Q.E.D.

The method in this section is called NMSEMS in De Leeuw (1968, 1970). It is closely related to the procedure for finding an initial configuration in the GL-SSA programs (see Guttman 1968). The NMSEMS solution is indeed extremely fast compared with other NMS-methods. I do not know how dirty it is. At the moment I am inclined to think that it is theoretically somewhat less satisfactory than the projection-type algorithms (Kruskal's MDSCAL, GL-SSA, Roskam's MINISSA, my own NMSPOM, Young's TORSCA, and so on). But I am also inclined to think that it may (in general) give more satisfactory numerical results, mainly because I distrust the iterative 'improvements' in the projection-type algorithms. All this is very hypothetical of course

2 Existence and uniqueness

2.1 The set of solutions

The inequalities (1.1) can be rewritten as

$$\sigma_{ijkl}\ \mathrm{Tr}\left\{X'(A^{ij} - A^{kl})X\right\} \geqslant 0.$$

We can drastically reduce the number of inequalities in this system by deleting the inessential ones (that follow by transitivity, reflexivity, or antisymmetry from other inequalities). We may also delete the inequalit that are automatically true for all X (such as $d_{kl} \geqslant d_{ii}$). In this reduce system we have, say, m inequalities of the form

$$\mathrm{Tr}(X'B_kX) \geqslant 0 \qquad (k=1,\ldots,m) \tag{1}$$

and we want to find n x n matrices X that satisfy these m constraints. Observe that in this form one can require either semi-strong or weak order-isometricity. The p-solutions for p=1,...,n-1 , if they exist, are among the n-solutions. More precisely: we redefine p-solutions as n-soluti of rank p.

Let S stand for the set of solutions to the inequalities (2). If X is a solution, then so is λX for all real λ. It follows that S is a bundle of lines through the origin (of the n^2-dimensional real linear space of all real square matrices of order n). Suppose that X and Y are two solutions. Then considering

$$\mathrm{Tr}\left\{\left[\lambda X + (1-\lambda)Y\right]' B_k \left[\lambda X + (1-\lambda)Y\right]\right\} =$$
$$= \lambda^2 \mathrm{Tr}(X'B_kX) + (1-\lambda)^2 \mathrm{Tr}(Y'B_kY) + 2\lambda(1-\lambda)\mathrm{Tr}(X'B_kY)$$

shows that S is not convex in general.

Translations are always permitted. Algebraically: if D is an arbitrary diagonal matrix, then ED is a solution (E is defined as in section 1). For all $X \in S$ it is true that the subspace spanned by X and ED is a subset of S. Rotations are always permitted. If K is a square orthonormal matrix, then for all $X \in S$ it is true that $XK \in S$. We can remove these indetermina cies by requiring in addition that the columns of X must be both centered and pairwise orthogonal, but it is difficult to see the effect this has o

2.2 Convexifying the problem

By using the identity

$$\mathrm{Tr}(X'B_kX) = \mathrm{Tr}(B_kXX')$$

we may again reformulate the problem. Find an n x n real matrix C such that

$$\mathrm{Tr}(B_kC) \geqslant 0 \qquad k=1,\ldots,m. \qquad (2)$$

This is a set of linear inequalities in (the elements) of C.

Lemma 2.1: C can be decomposed as $C = XX'$ with X real n x n iff C is symmetric PSD.

Proof: Necessity: If $C = XX'$ with X real, then $y'Cy = y'XX'y = z'z$ (say) $\geqslant$ for all real y. Thus C is PSD. Sufficiency: We simply mention three possibl decompositions. K are the eigenvectors of C (without loss of generality $K'K = KK' = I$), and Λ are the corresponding eigenvalues (collected in a diagonal matrix, all diagonal elements nonnegative). Then $C = K\Lambda K'$, choose $X = K\Lambda^{\frac{1}{2}}$. (orthogonal decomposition). Alternatively: choose $X = K\Lambda^{\frac{1}{2}}K'$ (symmetric decomposition). Alternatively: the Cholesky process proves that X can be chosen in lower-diagonal form (triangular decomposition). Q.E.D.

Lemma 2.2: The columns of X are centered iff C is doubly centered.

Proof: If $e'X = 0$ then $e'C = e'XX' = 0$, and $Ce = XX'e = 0$. Consequently C is doubly centered. For all real X it is true that $XX'y = 0$ implies $X'y = 0$. Thus if $Ce = XX'e = 0$, then $X'e = 0$, which proves the converse. Q.E.D.

From now on we work in the $\frac{1}{2}n(n-1)$ dimensional real linear space of all symmetric doubly centered (SDC) matrices of order n. The m linear inequaliti (2) define a polyhedral convex cone in this space.

D2

Lemma 2.3: The set T of all PSD matrices is a (nonpolyhedral) convex cone.

Proof: If C is PSD, then λC with $\lambda \geqslant 0$ is PSD. If C_1 and C_2 are PSD, then so is $\lambda C_1 + (1-\lambda)C_2$ for all $0 \leqslant \lambda \leqslant 1$. Using the fact that a necessary and sufficient condition for a real symmetric matrix C to be PSD is that $\mathrm{Tr}(CB) \geqslant 0$ for all PSD B, it follows that T is the intersection of an infinite number of halfspaces, i.e. T is not polyhedral. Q.E.D.

Corrolary 2.1: There is no finite list of linear inequalities in C that is both necessary and sufficient for PSD.

We summarize our results in the following theorem.

Theorem 2.1: The solution set of the following system of inequalities and equations

$$\left.\begin{array}{ll} \mathrm{Tr}(B_k C) \geqslant 0, & k=1,\ldots,m \\ \mathrm{Tr}(E^{i\cdot} C) = 0, & i=1,\ldots,n \\ C = C', & \\ C \text{ PSD}, & \end{array}\right\} C \quad \text{SDC}$$

is a nonpolyhedral convex cone in a linear space of $\frac{1}{2}n(n-1)$ dimensions.

Theorem 2.2: If $C = XX'$ and $C = YY'$ then there is a rotation matrix K (with $KK' = K'K = I$) such that $X = YK$. Conversely, if K is a rotation matrix, $C = YY'$ and $X = YK$, then $C = XX'$.

Proof: $XX' = YKK'Y' = YY' = C$, which proves the second part. If $C = XX'$ and $C = YY'$, then X can be written (uniquely) as $M_1\Lambda_1L_1'$ and Y as $M_2\Lambda_2L_2'$ (with the M_i and L_i square orthonormal, and Λ_i diagonal). It follows that $C = M_1\Lambda_1^2M_1' = M_2\Lambda_2^2M_2'$, and thus that $M_1 = M_2 = M$ (say), $\Lambda_1 = \Lambda_2 = \Lambda$(say). Let $K = L_2L_1'$, then K is square orthonormal and $YK = M\Lambda L_2'L_2L_1' = M\Lambda L_1' = X$. Q.E.D.

Thus we have eliminated the indeterminacy due to rotation and reflexion in our original formulation of the problem. The indeterminacy due to translations was already eliminated. Only uniform stretching and shrinking remains permitted (i.e. λC with $\lambda \geqslant 0$ instead of C), but this can be eliminated by requiring $\|C\| = 1$, where $\| \quad \|$ is any matrix norm. Observe that adding $\|C\| \leqslant 1$ to the requirements makes the solution set into a closed and bounded (and thus compact) convex set.

The structure of the cone P of solutions can be described more precisely. Let P_r denote the subset of all solutions with rank r, $r=0,\ldots,n-1$. Then $P_0, P_1, \ldots, P_{n-1}$ is a partitioning of P. The sets P_r ($r \geqslant 1$) are bundles of of half-open rays: if C is of rank r, then so is λC with $\lambda > 0$. The set P_0 is a singleton, only $0 \in P_0$. The P_r are most certainly not convex. If $C_1, C_2 \in P_r$ then in most cases $\mu C_1 + (1 - \mu)C_2 \in P_s$ with $s > r$. In fact $r \leqslant s \leqslant \min(2r, n - 1)$. A more elegant way to partition the cone P is made possible by the following theorem.

<u>Theorem 2.3</u>: If A and B are PSD matrices with the same null-space (and consequently the same rank), then $A + B$ also has the same null-space (and thus the same rank).

Proof: Suppose A and B have the same null-space. If $x'Ax = 0$ then $x'Bx = 0$ and $x'(A + B)x = 0$. The null-space of $A + B$ is a subspace of that of A (a that of B). If $x'(A + B)x = 0$ then $x'Ax + x'Bx = 0$, and because both A and are PSD, $x'Ax = x'Bx = 0$. The null-spaces of $A + B$ and A (and B) are thus identical, and rank $(A + B)$ = rank (A) = rank (B). Q.E.D.

The relation of having the same null-space is an equivalence relation, th partitions P into equivalence classes $P/\approx$.

<u>Corrolary 2.2</u>: The equivalence classes $P/\approx$ are convex cones.

2.3 <u>Representation theorem</u>

Another basic result is the following existence theorem.

D12 $\rightarrow$ <u>Theorem 2.4</u>: If ϕ is a POG-distance, then there always exists a strong $(n-1)$-representation. Moreover this representation can be chosen in such a way that the columns of X are orthogonal and that either the endpoints o all n vectors x_i lie on the unit sphere, or that the centroid of the configuration is the origin.

Proof: We give a constructive proof. First we embed our partial order in weak order. This can of course be done, but the embedding is not unique. Then we form the matrix T with all diagonal elements equal to zero, the off-diagonal element corresponding with the smallest δ_{ij} equal to -1, t next equal to -2, and so on (ties are averaged). If $\delta_{ij} \gtrsim \delta_{kl}$ then there is a requirement

$$\mathrm{Tr}\left\{(A^{ij} - A^{kl})T\right\} > 0,$$

or $t_{ii} + t_{jj} - t_{ij} - t_{ji} - t_{kk} - t_{ll} + t_{kl} + t_{lk} > 0$. Now $t_{ii} = t_{jj} = t_{kk} = t_{ll} = 0$, $t_{ij} = t_{ji}$, $t_{kl} = t_{lk}$, $t_{ij} < t_{kl}$ iff $\delta_{ij} \geq \delta_{kl}$, and thus this strict inequality is met. The same thing is true at the places where we require equality. Thus T fulfills all inequalities and equalities strictly, but it is not PSD. The form of the matrices A^{ij}, however, guarantees that we may add an arbitrary diagonal matrix to T, and still maintain strict monotonicity. We do this by determining the smallest eigenvalue of T (which is negative of course), call it λ , and define

$$R = -\frac{1}{\lambda}(T - \lambda I).$$

Then R is still strictly monotone, but it also is PSD. Its rank is n-1 (or less when λ is a multiple eigenvalue). Moreover all x_i do have unit lenght. The strict (n-1)-solution is given by the n - 1 scaled eigenvectors corresponding with the nonvanishing eigenvalues. The second type of solution mentioned in the theorem is obtained by translating the centroid to the origin and rotating to principal axes. Q.E.D.

D3 $\longrightarrow$ <u>Corollary 2.3</u>: If ϕ is a PCG-distance and there is at least one pair of dissimilarities such that $\delta_{ij} \geq \delta_{kl}$, then the cone of theorem 2.1 has a nonvoid interior.

2.4 <u>The uniqueness question</u>

Theorem 2.4 gives us a solution to the representation problem, theorem 2.1 says something about the uniqueness. Not much, however. The solution set is a convex cone with (except in trivial cases) a nonempty interior. We want to find a measure that expresses the degree of uniqueness of a solution, or, more precisely, we want a measure that says how large this cone is. It is intuitively clear, that for this we must investigate the boundary of the cone. Our further procedures will use the fact that a nonpolyhedral convex cone can be approximated to an arbitrary close degree by polyhedral convex cones (both from the inside and from the outside). Because the cone of all PSD-matrices id pointed, our approximating cone will be pointed, and we can speak of the edges of the (approximating) cone. We need some preliminary definitions and theorems. Define

$$r(C_1,C_2) = \frac{\mathrm{Tr}(C_1C_2)}{\sqrt{\mathrm{Tr}(C_1^2)\ \mathrm{Tr}(C_2^2)}}.$$

If P is a cone of SDC-matrices then we use

$$r_P = \min_{C_1,C_2 \in P} r(C_1,C_2)$$

Dis as a measure of uniqueness. Trivially $-1 \leq r \leq +1$, and $r = +1$ iff the solutions are on a ratio scale. We prove some less obvious properties of r

<u>Theorem 2.5</u>: If C_1, C_2 are PSD, then $r(C_1,C_2) \geq 0$, and $r(C_1,C_2) = 0$ iff al vectors x are either in the null space of C_1 or in the null space of C_2.

Proof: Both C_1 and C_2 are PSD, i.e. can be written in canonical form $K\Lambda K'$, with K real orthonormal and Λ positive diagonal. Then $\mathrm{Tr}(C_1C_2) = \mathrm{Tr}(C_1K$ $= \mathrm{Tr}(K'C_1K\Lambda) = \mathrm{Tr}(\Lambda\,\mathrm{diag}(K'C_1K)) = \lambda_1 k_1'C_1k_1 + \lambda_2 k_2'C_1k_2 + \dots + \lambda_n k_n'C_1k$ ≥ 0. If L and ψ are the eigenvectors and eigenvalues of C_1, then also $\mathrm{Tr}(C_1C_2) = \psi_1 l_1'C_2l_1 + \dots + \psi_n l_n'C_2l_n \geq 0$. These sums vanish iff $k_i'C_1k_i =$ for all nonzero λ_i, and $l_i'C_2l_i = 0$ for all nonzero ψ_i. If Z_1 and Z_2 are the null spaces of C_1 and C_2, and Y_1 and Y_2 are the nonnull spaces, then t first condition reduces to $Y_2 \subseteq Z_1$ and the second to $Y_1 \subseteq Z_2$. Both conditi are equivalent to the assertion that the union of Z_1 and Z_2 exhausts the space. Q.E.D.

Another measure of relatedness that is used a great deal in psychometrics is the canonical correlation. Suppose that C_1 and C_2 are two doubly cente PSD matrices with unit Euclidean norm and canonical forms $K\Lambda K'$ and $L\Psi L'$. K has p_1 columns, L has p_2 columns, $X = K\Lambda^{\frac{1}{2}}$, $Y = L\Psi^{\frac{1}{2}}$. The symbol ρ_{ij} denotes the PM-correlation between column i of X and column j of Y. The $\min(p_1,p_2)$ canonical correlations are denoted by R_k.

<u>Theorem 2.6</u>:

$$\sum_k R_k^2 = \sum_i^{p_1}\sum_j^{p_2} \rho_{ij}^2,$$

$$r(C_1,C_2) = \sum_i^{p_1}\sum_j^{p_2} \lambda_i \psi_j \rho_{ij}^2.$$

Proof: $Tr(C_1C_2) = Tr(X'YY'X) = \sum_i \sum_j \lambda_i \psi_j \rho_{ij}^2$. The sum of the squared canonical correlations is equal to $Tr((X'X)^{-1}X'Y(Y'Y)^{-1}Y'X) = Tr(K'LL'K) = \sum_i \sum_j \rho_{ij}^2$. Q.E.D.

It follows that r is a sum of squared covariances, expressing the fact that differential stretching and shrinking of configuration matrices is not permitted. The close conceptual correspondence with canonical correlation is clear. Consider the problem: find rotation matrices M_1 and M_2 such that $Tr(M_2'Y'XM_1) = Tr(M_1'X'YM_2)$ is a maximum. The stationary equations are

$$Y'XM_1 = M_2D,$$
$$X'YM_2 = M_1D,$$

with D symmetric. It follows that $Tr(D^2) = Tr(X'YY'X) = r(C_1,C_2)$.

Corollary 2.4: If x_1 and x_2 are two centered one-dimensional solutions, then

$$r(C_1,C_2) = \rho^2(x_1,x_2),$$

where ρ denotes the PM correlation coefficient.

Corollary 2.5: If the MDS-problem has one-dimensional solutions x and y then

$$r_{min}(C_1,C_2) \leq \rho^2_{min}(x,y),$$

where the minimum is taken in the first case over all (n-1)-solutions, in the second case over all 1-solutions.

These two corollaries give the connection with previous work in the one-dimensional field (Abelson and Tukey (1963,1959), Lindman (1969)). The next theorem is our main result on uniqueness.

Theorem 2.7: If $C_1,C_2,\ldots,C_k$ are the edges of a pointed cone of SDC-matrices then there are indices i and j such that

$$\min_{A,B \in P} r(A,B) = r(C_i,C_j).$$

Proof: Without loss of generality we assume that the C_i have unit Euclidean norm. If $A,B \in P$ then there are k nonnegative numbers α_i such that $A = \sum \alpha_i C_i$, and k nonnegative numbers β_i such that $B = \sum \beta_i C_i$. Therefor

$$Tr(AB) = \sum \sum \alpha_i \beta_j Tr(C_iC_j) = \sum \sum \alpha_i \beta_j r(C_i,C_j) \geq$$

$$\geqslant \sum \alpha_i \sum \beta_i \min_{i,j} r(C_i, C_j).$$

Moreover

$$\sqrt{\mathrm{Tr}(A^2)} = \sqrt{\sum \sum \alpha_i \alpha_j r(C_i, C_j)} \leqslant \sqrt{(\sum \alpha_i)^2 \max_{i,j} r(C_i, C_j)} = \sum$$

In the same way

$$\sqrt{\mathrm{Tr}(B^2)} \leqslant \sum \beta_i.$$

Combining these results yields

$$r(A,B) \geqslant \min_{i,j} r(C_i, C_j).$$

Q.E.D.

It follows that if we want to compute r_P for a particular cone P, it suffi to compute the edges and their intercorrelations. For a polyhedral cone th number of edges is finite and this procedure can be carried out. Neverthel it may be a formidable task, even for the fastest electronic computers.

Our coefficient r is related to the Euclidean norm and to the Euclidean distance between the endpoints of unit lenght vectors. A different norm is sometimes also useful.

Theorem 2.8: The sum of the diagonal elements is an additive norm over the cone of all PSD matrices.

Proof: If C is PSD, then $c_{ii} \geqslant 0$ for all i, and consecuently $\mathrm{Tr}(C) \geqslant 0$. Moreover C is PSD and $c_{ii} = 0$ for all i iff $C = 0$. Tr(C) is linear, which implies that it is additive (and a fortiori subadditive). Finally we must prove that $\mathrm{Tr}(CD) \leqslant \mathrm{Tr}(C)\mathrm{Tr}(D)$ for all PSD C and D. We use the canonical form of C as in the proof of theorem 2.5. $\mathrm{Tr}(CD) = \sum \lambda_i k_i' D k_i \leqslant \sum \lambda_i \sum k_i' D k_i$ $\leqslant \mathrm{Tr}(C)\mathrm{Tr}(D)$. Q.E.D.

Adding the requirement $\mathrm{Tr}(C) = n-1$ to the ones mentioned in theorem 2.1 has the effect that we eliminate the indeterminacy due to uniform stretch and shrinking, that our solution set becomes a closed and bounded (and th compact) convex set, and that the approximating cones are replaced by closed and bounded convex polyhedra. For our 'extreme' solutions we only have to consider the vertices of these polyhedra.

3 Linearizing the problem

3.1 Sufficient conditions

We consider again the inequalities and equations

$$\begin{aligned} Tr(B_k C) &\geqslant 0, \qquad (k=1,\ldots,m) \\ C &\quad SDC, \\ C &\quad PSD. \end{aligned}$$

Call the solution set (a nonpolyhedral convex cone) P. Corollary 2.1 state that 'C PSD' can not be translated into an equivalent system of linear inequalities that is finite. We can, however, find finite systems of linear inequalities that are either sufficient or necessary. We start with the sufficient conditions.

Theorem 3.1: If $Tr(D_l C) \geqslant 0 \quad (l=1,\ldots,p)$ is a set of conditions sufficie for PSD of C, and $\tilde{P}$ is the solution set of

$$\begin{aligned} Tr(B_k C) &\geqslant 0, \\ C &\quad SDC, \\ Tr(D_l C) &\geqslant 0, \end{aligned}$$

then $\tilde{P}$ is a polyhedral convex cone and $\tilde{P} \subset P$.

Proof: Rather obvious. If $C \in \tilde{P}$, then C is PSD, and thus $C \in P$. Corollary 2.1 implies that there is at least one PSD $C \notin \tilde{P}$. Q.E.D.

Corollary 3.1:

$$\min_{C_1, C_2 \in \tilde{P}} r(C_1, C_2) \geqslant \min_{C_1, C_2 \in P} r(C_1, C_2).$$

The most practical set of linear inequalities sufficient for PSD that I have be able to come up with is

$$\begin{aligned} c_{ij} &\leqslant 0, \qquad (\forall i \neq j) \\ \sum_j c_{ij} &\geqslant 0. \qquad (\forall i) \end{aligned}$$

The last set of n inequalities is redundant, because it is implied by SDC. Remembering definition 1.1, it is clear that we want C to be a doubly centered T-matrix (condition DCT). The system is

$$\mathrm{Tr}(B_k C) \geqslant 0, \tag{1a}$$

$$C \quad \mathrm{DCT}. \tag{1b}$$

3.2 The T-method.

This suggest another quick-and-dirty method: study the cone $\tilde{P}$ defined by the requirements (3.1.1).

Lemma 3.1: The requirement that C must be a T-matrix defines a pointed polyhedral convex cone in the $\frac{1}{2}n(n+1)$-dimensional linear space of all real symmetric matrices of order n. The edges of this cone are the $\frac{1}{2}n(n-1)$ matrices A^{ij} and the n matrices E^{ii} defined in section 1. The requirement that C must be a doubly centered T-matrix defines another pointed cone P_D in the $\frac{1}{2}n(n-1)$-dimensional space of all symmetric DC-matrices with the $\frac{1}{2}n(n-1)$ edges A^{ij}.

D13 $\longrightarrow$ Proof: The requirement 'C T' is equivalent to the $\frac{1}{2}n(n+1)$ linear inequalities

$$c_{ij} \leqslant 0, \qquad (\forall i \neq j),$$
$$\sum_j c_{ij} \geqslant 0, \qquad (\forall i).$$

All inequalities are satisfied as equations iff $C = 0$, thus the cone is pointed. All inequalities but one from the first set are satisfied as equations by the A^{ij}. And all but one from the second set by the E^{ii}. These are the edges. The cone P_D is spanned by the A^{ij} because if

$$C = \sum_{i} \sum_{\neq j} \lambda_{ij} A^{ij} + \sum_{i=1} \beta_i E^{ii},$$

then C is DC iff $\beta_i = 0$ for all $i \in N$, and thus the A^{ij} are the edges. Q.E.

Observe that for this cone P_D we already have $r_{min} \geqslant 0$. In fact it is equal to zero if we choose A^{ij} and A^{kl} with $i \neq j \neq k \neq l$. If $n = 3$ this can not be done, and $r_{min} = 0.25$. Because the cone $\tilde{P}$ defined by (3.1.1) is a subcone of P_D, we have of course

$$\min_{C_1, C_2 \in \tilde{P}} r(C_1, C_2) \geqslant \min_{C_1, C_2 \in P_D} r(C_1, C_2).$$

In the next theorem we shall sketch our most important computational procedure. If $Q = \left\{ C \mid \mathrm{Tr}(B_k C) \geq 0 \ ; \ k=1,\ldots,m \right\}$ is a pointed cone with edges $\left\{ C_1,\ldots,C_l \right\}$ in the $N = \frac{1}{2}n(n-1)$ dimensional space of all real SDC-matrices of order n, and $\overline{Q} = Q \cap \left\{ C \mid \mathrm{Tr}(B_{m+1}C) \geq 0 \right\}$, then form the l-element vector $d_i = \mathrm{Tr}(B_{m+1}C_i)$. If d has n_0 zero elements, n_+ positive and n_- negative elements, then form a set E of $n_0 + n_+ + n_+n_-$ SDC-matrices of order n by the rules

$$\text{i)} \quad d_i \geq 0 \Rightarrow C_i \in E,$$

$$\text{ii)} \quad d_i > 0 \wedge d_j < 0 \Rightarrow d_i C_j - d_j C_i \in E.$$

<u>Theorem 3.2</u>: The matrices in E span the cone $\overline{Q}$ (i.e. the edges of $\overline{Q}$ are elements of E).

Proof: Of course $\overline{Q}$ is also pointed. C is an edge of $\overline{Q}$ iff $\mathrm{Tr}(B_k C) \geq 0$ with equality for N - 1 linear independent matrices B_k ($k = 1,\ldots,m+1$). There are two possibilities. If the N - 1 equations $\mathrm{Tr}(B_k C) = 0$ are all satisfied for $k \leq m$, then C is also an edge of Q, and $\mathrm{Tr}(B_{m+1}C) \geq 0$, so the first rule takes care of these edges. If there are only N - 2 equations for $k \leq m$, then C must lie in one of the two dimensional faces of C. In order to be an edge of $\overline{Q}$ it must also lie in the hyperplane $\mathrm{Tr}(B_{m+1}C) = 0$. The two dimensional faces of Q are among the sets $\left\{ C \mid C = \alpha C_i + \beta C_j \ ; \ \alpha,\beta \geq \right.$ Intersecting this set with the hyperplane produces a C matrix which is not already included among the elements of E by application of the first rule iff $d_i > 0$ and $d_j < 0$ or vice versa. The second rule takes care of these edges. Q.E.D.

This procedure is very similar to procedures used by Uzawa (1958) and Lindman (1969). The difference is that we operate with one added inequality at a time (which has advantages in our later procedures) and that we start with a pointed cone (which can be done because of our particular problem). The final assumption makes our proof much simpler than the ones given by Uzawa and Lindman (the theorem is of course less general).

The procedure will be illustrated by an example in which $n = 3$, and δ_{12}
$\delta_{13} \gtrsim \delta_{23}$. We start with the three edges of the DCT-cone P_D.

$$\begin{array}{ccc} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \\ & (1) & \end{array} \qquad \begin{array}{ccc} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \\ & (2) & \end{array} \qquad \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \\ & (3) & \end{array}$$

The value of all three r_{ij} is equal to 0.25. Our first inequality demands
that $d_{12}^2 - d_{13}^2 \geqslant 0$, or $Tr(B_1C) \geqslant 0$, with B_1 =

$$\begin{array}{ccc} 0 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & -1 \end{array}.$$

The three values of d_i are $(\ 3\ -3\ \ 0)$, which means that the new spanning
set E contains numbers (1) and (3) (by rule i), and the sum of (1) and (2
(by rule ii).

$$\begin{array}{ccc} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \\ & (1) & \end{array} \qquad \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \\ & (2) & \end{array} \qquad \begin{array}{ccc} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ & (3) & \end{array}$$

Now $r_{12} = 0.25$, $r_{13} \approx .79$, $r_{23} \approx .32$, which means that r_{min} is still equa
to 0.25. The next inequality is $d_{13}^2 \geqslant d_{23}^2$, or $Tr(B_2C) \geqslant 0$, with B_2 =

$$\begin{array}{ccc} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{array}.$$

The values of d_i are $(\ 0\ -3\ \ 3)$, so the final set of edges (they are inde
all edges) is given by

$$\begin{array}{ccc} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \\ & (1) & \end{array} \qquad \begin{array}{ccc} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ & (2) & \end{array} \qquad \begin{array}{ccc} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \\ & (3) & \end{array}$$

Now $r_{12} \approx .79$, $r_{13} \approx .71$ (and minimal), $r_{23} \approx .89$. The corresponding
configurations are given in figure 1a,b,c.

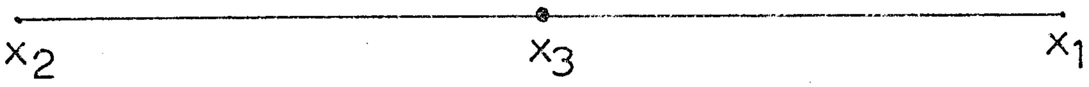


fig 1a

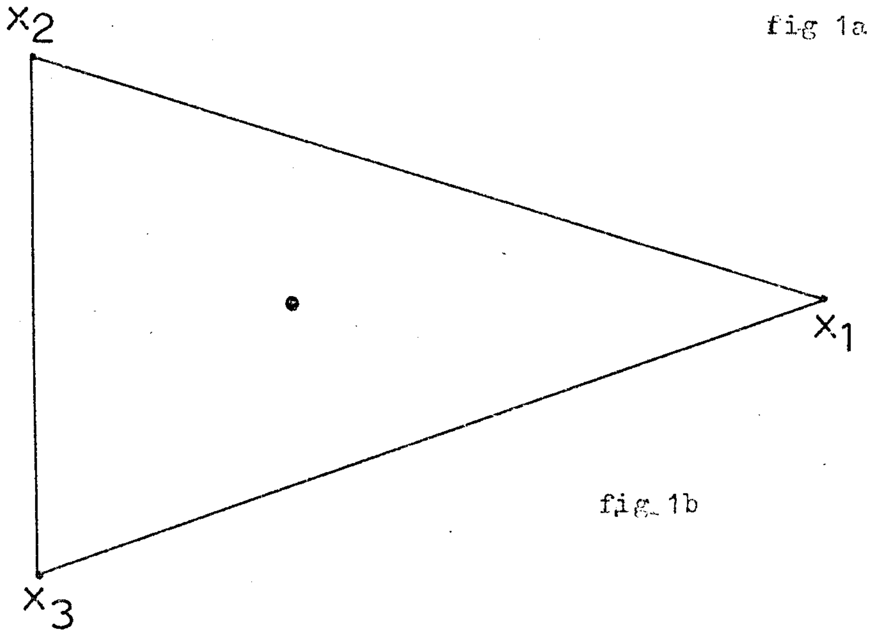


fig 1b

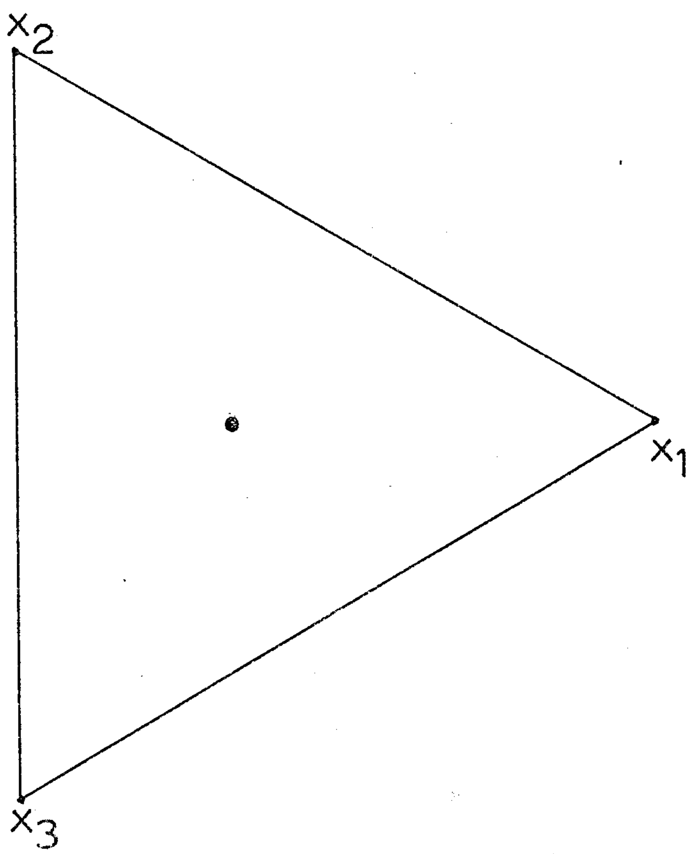
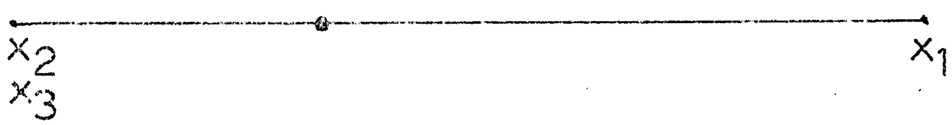


fig 1c

fig 1d



3.3 Necessary conditions

Theorem 3.3: If $Tr(D_l C) \geq 0$ $(l=1,\ldots,p)$ is a set of conditions necessary for PSD of C, and $\hat{P}$ is the solution set of

$$\begin{aligned} &Tr(B_k C) \geq 0 \\ &\quad C \quad SDC \\ &Tr(D_l C) \geq 0 \end{aligned}$$

then $\hat{P}$ is a polyhedral convex cone, and $P \subseteq \hat{P}$.

This theorem is the dual of theorem 3.1. The dual of corollary 3.1 follows.

Corollary 3.2:

$$\min_{C_1,C_2 \in \hat{P}} r(C_1,C_2) \leq \min_{C_1,C_2 \in P} r(C_1,C_2).$$

Let χ be the set of all real PSD symmetric matrices of rank one, i.e. the set of all matrices that can be written as $X = xx'$, with x a real vector.

Lemma 3.2: A matrix A is PSD iff $Tr(AY) \geq 0$ for all $Y \in \chi$.

Proof: If $Y \in \chi$ then there exists a real vector y such that $Y = yy'$. If A is PSD then $Tr(AY) = Tr(Ayy') = y'Ay \geq 0$ for all $Y \in \chi$. If $Tr(AY) \geq 0$ for all $Y \in \chi$, then $y'Ay \geq 0$ for all $y \in R^n$, and A is PSD. Q.E.D.

It follows that, given any vector y, the linear inequality $Tr(Cyy') \geq 0$ is necessary for PSD. Any finite list of these conditions is also necessary for PSD (but not sufficient).

3.4 The complete method

Our method can now be described completely. Suppose Q_s is a pointed polyhedral convex cone containing P. Moreover $C_1,\ldots,C_k$ is a spanning set for Q_s. Find the smallest eigenvalue λ_i of each of the C_i. Let C_s be the matrix for which λ_s is the smallest of the λ_i, and let y_s be the corresponding eigenvector of C_s. Then add the constraint $Tr(Cyy') \geq 0$. Compute the spanning set of the cone Q_{s+1} by the procedure outlined in theorem 3.2, normalize the elements of this set such that $Tr(C_i) = n - 1$, and start again.

D14

This procedure generates the sequences $\{Q_s\}$, a sequence of polyhedral convex cones, the sequence of matrices $\{C_s\}$, and the sequence of real nu bers $\{\lambda_s\}$. Moreover we compute in each iteration (using the elements in the spanning set)

$$r_s = \min_{C_1, C_2 \in Q_s} r(C_1, C_2),$$

and we obtain another sequence $\{r_s\}$.

D4

Theorem 3.4: This procedure generates an infinite sequence of nonempty polyhedral convex cones, that are related in the following way,

$$Q_0 \supset Q_1 \supset Q_2 \supset \dots$$

(and inclusion is strict). Moreover, all these cones contain P.

Proof: It is sufficient to prove that $\lambda_s < 0$ for all $s=0,1,2,\dots$ Suppose not. Then all elements in the spanning set are PSD, and consequently Q_s is a subset of the cone of all PSD matrices, which cannot be true because the conditions used in the definition of Q_s are only necessary (and not sufficient) for PSD. It follows that C_s violates the inequality $Tr(y_s y_s' C)$ 0, because $Tr(y_s y_s' C_s) = y_s' C_s y_s = \lambda_s < 0$. Thus $Q_{s+1} \subset Q_s$ (strictly). The conditions always remain only necessary for PSD, and thus $P \subset Q_s$ for all (also strictly). Q.E.D.

Corollary 3.3: The sequence $\{r_s\}$ is nondecreasing. Moreover $r_s \leq r_p$ for all s. The sequence $\{\lambda_s\}$ is increasing, $\lambda_s < 0$ for all s.

Corollary 3.4: The sequence $\{Q_s\}$ converges to a cone $\bar{Q} \supset P$. The sequence $\{r_s\}$ converges to a value $\bar{r} \leq r_p$. Finally $\lambda_s \to \bar{\lambda} \leq 0$.

In order to complete our convergence proof we reformulate the problem as

$$\lambda_{min}(C) \quad \min!$$

$$Tr(B_k C) \geq 0$$

$$C \quad SDC$$

$$\lambda_{min}(C) \geq 0.$$

This problem will be called problem (D).

Lemma 3.3: $\lambda_{min}(C)$ is a continuous concave matrix function with continuous matrix derivative

$$\frac{\partial \lambda_{min}(C)}{\partial C} = y_{min} y'_{min},$$

where y_{min} the the eigenvector corresponding with the smallest eigenvalue.

Proof: $\lambda_{min}(\mu C_1 + (1-\mu)C_2) = \mu y'_{min} C_1 y_{min} + (1-\mu) y'_{min} C_2 y_{min} \geqslant \mu \lambda_{min}(C_1) + (1-\mu)\lambda_{min}(C_2)$. The continuity of all roots is well known, the formula for the derivative follows by simple calculation (we make the simplifying assumption that the smallest root is not a multiple root). On a bounded set $\lambda_{min}(C)$ is bounded, and continuity follows from concavity. Q.E.D.

The scaling of solutions such that $Tr(C_i) = n - 1$ for all matrices in the spanning set means that we actually use this scaling requirement too. Interprete SDC from now on (and also in the formulation of problem D) as scaled, symmetric, and doubly centered. This implies, of course, that we are in fact working with polyhedra in stead of cones.

Theorem 3.5: The sequences $\{C_s\}$ and $\{\lambda_s\}$ produced by our procedure are the same as the sequences of solutions and function values obtained by applying the cutting plane procedure to problem (D).

Proof: The fact that in problem (D) the minimization subproblems are not linear but concave only complicates the computation, not the theoretical basis of the algorithm. It is still true that we find the solutions at the vertices of the permissible region of the subproblems. We have only one convex nonlinear restraint, $-\lambda_{min}(C) \leqslant 0$, the solution C_s of subproblem s violates this and has to be cut off. The added restriction which accomplishes this is (in the 'concave' cutting plane method)

$$-\lambda_{min}(C_s) - Tr\left[(C - C_s)\nabla\lambda_{min}(C_s)\right] =$$

$$-\lambda_s - y'_s C y_s + y'_s C_s y_s = -y'_s C y_s = -Tr(C y_s y'_s) \leqslant 0,$$

which is the same as in our method. Q.E.D.

The cutting plane method was discovered by Cheney and Goldstein (1959) and by Kelley (1960). Detailed descriptions can be found in Collatz and Wetterling (1966, p 93-98), and Zangwill (1969, chapter 14). To prove convergence we only have to apply the usual convergence theorems for the cutting plane method to our particular problem.

Theorem 3.6:

$$\text{i)}\quad \bar{\lambda} = 0,$$

$$\text{ii)}\quad \bar{r} = r_p,$$

$$\text{iii)}\quad \bar{Q} = P.$$

Proof: We investigate the sequence $\{C_s\}$. All C_s lie in the bounded set $Q_0 \cap \{C \mid Tr(C) = n - 1\}$. Therefore $\{C_s\}$ has a subsequence which converg to, say, $\bar{C}$. Suppose $\bar{C} \notin P$, then $\bar{\lambda} = \lambda_{min}(\bar{C}) < 0$. Because of the continuity of $\lambda_{min}(C)$ we know that there exists a C_t in the subsequence such that

$$\|C_t - \bar{C}\| < -\tfrac{1}{2}\bar{\lambda},$$

and

$$\lambda_{min}(C^t) > \tfrac{1}{2}\bar{\lambda}.$$

We now have the identity

$$\lambda_{min}(C_t) + Tr\left[(\bar{C} - C_t)y_t y_t'\right] = Tr(\bar{C}y_t y_t'),$$

as well as the inequality

$$\lambda_{min}(C_t) + Tr\left[(\bar{C} - C_t)y_t y_t'\right] < \tfrac{1}{2}\bar{\lambda} - \|\bar{C} - C_t\|\,\|y_t y_t'\| =$$

$$= \tfrac{1}{2}\bar{\lambda} - \|\bar{C} - C_t\| < 0.$$

It follows that $\bar{C} \notin Q_{t+1} = Q_t \cap \{C \mid Tr(Cy_t y_t') \geq 0\}$, which contradicts the fact that $\bar{C} \in \bar{Q} \subset Q_{t+1}$. Consequently $\bar{C} \in P$, and the theorem follows. Q.E.D.

We did not mention a stopping criterion in our description of the algorit The sequences $\{\lambda_s\}$ and $\{r_s\}$ are useful for this purpose. If we are minimizing λ (or: if we are only linearizing the constraints) then $\lambda_s >$ (where ε is some small positive number) seems the most useful criterion. If we are minimizing r (or: if we are interested in the uniqueness of the solutions) then we should stop as soon as the edges C_i and C_j of Q_s for which $r_s = r(C_i,C_j)$ are PSD (finite number of steps) or as soon as $r_{s+1} -$ $< \varepsilon$. That the sequence $\{r_s\}$ can indeed converge to r_p in a finite numb of steps is shown in some of the examples later on. The precise conditions under which this may happen are not known. Using some results of Wolfe (19 it can be shown that if the r_{min} given by the T-method of section 3.2 is equal to r_p (i.e. if the matrices $C_1,C_2 \in P$ such that $r_p = r(C_1,C_2)$ are T-matrices) then the procedure can be finite. Another example shows that in some cases there may exist a finite number of conditions $Tr(Cyy') \geq 0$ which, taken together with the requirements $Tr(B_kC) \geq 0$, are sufficient for PSD. In that case finite convergence can occur too. The influence of Wolfe' accelerating device was not tried out on our problem.

3.5 Computational remarks

Of course we can use a number of cutting planes in each step. All vectors y for which $y'C_iy < 0$ for some i will do. It seems easier, however, to proceed with one inequality at a time. The minimum eigenvalue and the corresponding eigenvector do not have to be computed with great precision. As long as (the Rayleigh quotient) $\lambda_s < 0$ we cut off something, and convergence is assured. If we are only minimizing r, then we do not have to compute the smallest eigenvalue of all C_i, only of those two for which $r(C_i,C_j) = r_s$ (and this may be important because we know that r_s may conver in a finite number of steps). These dilemma's are of course well known from other kinds of iterative procedures: with a larger amount of work in a step we need less steps.

Observe that our problem consists (computationally) of two different stage The first stage is to find an initial cone Q_0. We want that cone to be a good approximation of P, and we want the edges of Q_0 to be easy to compute. This is exactly analogous to finding a good starting value in other kinds o iterative programs. The second stage is to approximate P as quickly and efficiently as possible. In this respect the most obvious disadvantage of our complete method is that the number of elements in the spanning set may multiply at an alarming rate. In a rather optimistic case this number will obey the difference equation $n_{s+1} = 2(n_s - 1)$. Using the initial condition $n_0 = \binom{n}{2}$, we have after s steps,

$$n_s = 2^s \left[\binom{n}{2} - 2 \right] + 2,$$

and we must store $(\binom{n}{2} + n)n_s$ numbers. For $n = 10$, $s = 10$ this is already larger than 10^6. Therefore,it could be important to delete the elements of the new spanning set that are not edges (i.e. those which are positive linear combinations of other elements in the spanning set). This can be done by investigating the number of inequalities satisfied as equations, or by solving the sequence of LP-problems

$$\sum_{i \leqslant} \sum_{j} (c^k_{ij} - \sum_{p \neq k} q_p c^p_{ij}) \quad \min !$$

$$c^k_{ij} - \sum_{p \neq k} q_p c^p_{ij} \geqslant 0 \qquad \forall i \leqslant j$$

$$q_p \geqslant 0 \qquad \forall p$$

for each C_k in the spanning set. If the minimum value of this problem is zero, then delete C_k and select one of the remaining C_i for a new test. This means, of course, a considerable reduction of memory requirements, but a considerable increase in computer time.

Despite all the possible refinements in our method, it will remain a curiosity which can only be applied to very small examples, and which is most certainly not a candidate for routine application. An algorithm which solves the same problem in a somewhat different way will be discussed in a later section. This algorithm is expected to be faster, and it requires only a relatively small amount of storage. Nevertheless it is still quite expensive and its use in the future may still be limited to investigating certain theoretical problems (such as uniqueness) or to attacking some particularly nasty cases of degeneracy. The algorithm in section 3.4 can be compared with Uzawa's LP-method (1958b), or with the also closely related double description method (Motzkin et al 1953). These methods also solve the LP-problem by constructing and investigating all basic feasible solutions, which means that they do give considerably more information than the simplex method. Because you cannot get something from nothing they are much less efficient.

3.6 Initial spanning set

Suppose that all 'upper-diagonal' dissimilarities are weakly ordered and that there are no ties. In this case we have $\frac{1}{2}n(n-1)-1$ inequalities

$$\mathrm{Tr}(B_k C) \geqslant 0.$$

We add the n inequalities (necessary conditions)

$$\mathrm{Tr}(E^{ii} C) \geqslant 0,$$

and the single inequality (necessary condition)

$$\mathrm{Tr}(EC) \geqslant 0.$$

This makes a total of $\frac{1}{2}n(n+1)$ inequalities and the cone turns out to be pointed. The edges can be found relatively easily. The final inequality corresponds with the edge $C_1 = E - I$. The middle n inequalities with the n edges defined by

$$C_i = \frac{1}{n-1} I - \frac{1}{n-1} E + \tfrac{1}{2}E^{i\cdot} + \tfrac{1}{2}E^{\cdot i}.$$

The first $\frac{1}{2}n(n-1)$ inequalities correspond with equally many edges defined by

$$C_{ij} = \frac{-2(\rho_{ij}-1)}{n(n-1)} H_{ij} + \frac{n(n-1)-2(\rho_{ij}-1)}{n(n-1)} G_{ij},$$

with ρ_{ij} defined as the rank number of δ_{ij},

$$H_{ij} = \{h^{ij}_{kl}\} = \begin{cases} 1 & \text{if } \delta_{kl} > \delta_{ij}, \\ 0 & \text{otherwise.} \end{cases}$$

$$G_{ij} = \{g^{ij}_{kl}\} = \begin{cases} 1 & \text{if } \delta_{kl} < \delta_{ij}, \\ 0 & \text{otherwise.} \end{cases}$$

Thus $G_{ij} = E - I - H_{ij}$. The next thing to do is to project the cone into the subspace with $\mathrm{Tr}(E^{\cdot i}C) = 0$ for all $i \in N$. This is done by a procedure which is very similar to the one explained in theorem 3.2. We have $p = \frac{1}{2}n(n+1)$

edges $C_1,\ldots,C_p$. Compute $d_i = \mathrm{Tr}(E^{\cdot 1}C_i)$. The numbers n_0, n_-, and n_+ are defined as usual, but now the new spanning set contains $n_0 + n_- n_+$ elements (we retain only those edges for which $d_i = 0$). Then compute the new d_i, and so on. We end up (in this case) with $\frac{1}{2}n(n - 1)$ edges (i.e. n less than we started with). In fact it is easy to prove that this procedure is a special case of the one outlined previously. Apply $\mathrm{Tr}(E^{\cdot i}C) \geqslant 0$ first, then apply $-\mathrm{Tr}(E^{\cdot i}C) \geqslant 0$.

Example: again $\delta_{12} \geqslant_0 \delta_{13} \geqslant_0 \delta_{23}$. The initial six edges are

$$\begin{array}{ccc}
\begin{matrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix} & \begin{matrix} 2 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{matrix} & \begin{matrix} 0 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 0 \end{matrix} \\
(1) & (2) & (3) \\
\begin{matrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{matrix} & \begin{matrix} 0 & -2 & 1 \\ -2 & 0 & 1 \\ 1 & 1 & 0 \end{matrix} & \begin{matrix} 0 & -1 & -1 \\ -1 & 0 & 2 \\ -1 & 2 & 0 \end{matrix} \\
(4) & (5) & (6)
\end{array}$$

The vector d is (2 2 -1 -1 -1 -2), which means that there are eight matrices in the next set.

$$\begin{array}{cccc}
\begin{matrix} 0 & 1 & -1 \\ 1 & 4 & 1 \\ -1 & 1 & 0 \end{matrix} & \begin{matrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 4 \end{matrix} & \begin{matrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix} & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{matrix} \\
(1) & (2) & (3) & (4) \\
\begin{matrix} 2 & 0 & -2 \\ 0 & 4 & -1 \\ -2 & -1 & 0 \end{matrix} & \begin{matrix} 2 & -2 & 0 \\ -2 & 0 & -1 \\ 0 & -1 & 4 \end{matrix} & \begin{matrix} 2 & -4 & 2 \\ -4 & 0 & 1 \\ 2 & 1 & 0 \end{matrix} & \begin{matrix} 2 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{matrix} \\
(5) & (6) & (7) & (8)
\end{array}$$

The origin of these new elements, and their next d_i-value (sum of the second row) is listed in the following table.

number	origin	d_i
(1)	(1) + 2 x (3)	6
(2)	(1) + 2 x (4)	0
(3)	(1) + 2 x (5)	0
(4)	(1) + (6)	1
(5)	(2) + 2 x (3)	3

(6)	(2) + 2 x (4)	-3
(7)	(2) + 2 x (5)	-3
(8)	(2) + (6)	0

This gives us nine elements in the new spanning set.

$$\underset{(1)}{\begin{matrix} 4 & -3 & -1 \\ -3 & 4 & -1 \\ -1 & -1 & 8 \end{matrix}} \qquad \underset{(2)}{\begin{matrix} 4 & -7 & 3 \\ -7 & 4 & 3 \\ 3 & 3 & 0 \end{matrix}} \qquad \underset{(3)}{\begin{matrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 4 \end{matrix}} \qquad \underset{(4)}{\begin{matrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix}}$$

$$\underset{(5)}{\begin{matrix} 2 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & 4 \end{matrix}} \qquad \underset{(6)}{\begin{matrix} 2 & -4 & 2 \\ -4 & 0 & 4 \\ 2 & 4 & 0 \end{matrix}} \qquad \underset{(7)}{\begin{matrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{matrix}} \qquad \underset{(8)}{\begin{matrix} 4 & -4 & 0 \\ -4 & 4 & 0 \\ 0 & 0 & 0 \end{matrix}}$$

$$\underset{(9)}{\begin{matrix} 2 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{matrix}}$$

(1)	(1) + 2 x (6)	6
(2)	(1) + 2 x (7)	6
(3)	(2)	6
(4)	(3)	2
(5)	3 x (4) + (6)	6
(6)	3 x (4) + (7)	6
(7)	(5) + (6)	0
(8)	(5) + (7)	0
(9)	(8)	0

The final spanning set (i.e. the edges of Q_0) is given by the three matric

$$\underset{(1)}{\begin{matrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{matrix}} \qquad \underset{(2)}{\begin{matrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{matrix}} \qquad \underset{(3)}{\begin{matrix} 2 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{matrix}}$$

R:				Spectra:			
(1)	1.000	.707	.447	(1)	3	3	0
(2)		1.000	.632	(2)	6	0	0
(3)			1.000	(3)	9	0	-3

The spectra are scaled, the values of r are rounded. The eigenvector associated with the eigenvalue -3 is (0 1 -1), which means that we must add the inequality $\mathrm{Tr}(CD) \geqslant 0$, with $D = yy' =$

$$\begin{matrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{matrix}$$

The new d_i values are (6 1 −2), which gives us the four edges

$$\begin{matrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{matrix} \qquad \begin{matrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{matrix} \qquad \begin{matrix} 4 & -2 & -2 \\ -2 & 1 & 1 \\ -2 & 1 & 1 \end{matrix} \qquad \begin{matrix} 4 & -3 & -1 \\ -3 & 2 & 1 \\ -1 & 1 & 0 \end{matrix}$$

(1) (2) (3) (4)

R:	(1)	1.000	.707	.707	.655	Spectra:	(1)	3	3	0
	(2)		1.000	.750	.926		(2)	6	0	0
	(3)			1.000	.926		(3)	6	0	0
	(4)				1.000		(4)	6.4	0	−.4

We know now that $r_p \geqslant .655$. We also know (from the T-method of section 3.2 that $r_p \leqslant .707$. The same conclusion follows from the example in this secti because $r_{12} = .707$ and both (1) and (2) are PSD. The configuration corres- ponding with edge (3) is given in figure 1d. Observe that if we had requir in addition that $d^2_{12} \leqslant d^2_{13}$, then there would have been only two edges, (1) and (3), both PSD, and $r_p = .707$.

An alternative procedure (which can also be used in the case of partially ordered data) is to start with the necessary conditions

$$c_{ii} \geqslant 0 \qquad \forall i \in N,$$

$$c_{ii} + c_{jj} - 2c_{ij} \geqslant 0 \qquad \forall i \neq j \in N.$$

These inequalities define a pointed cone with $\frac{1}{2}n(n+1)$ edges

$$E^{i\cdot} + E^{\cdot i} \qquad \forall i \in N,$$

$$-(E^{ij} + E^{ji}) \qquad \forall i \neq j \in N.$$

The minimum value of r for this cone is already $-\frac{1}{n-1}\sqrt{n-1}$. Then we apply the equations

$$\sum_j c_{ij} = 0 \qquad \forall i \in N,$$

and the inequalities

$$\mathrm{Tr}(B_k C) \geqslant 0 \qquad k=1,\ldots,m,$$

which gives us Q_0.

We can also start with the system

$$\sum_j c_{ij} = 0 \qquad \forall i \in N,$$

$$c_{ii} + c_{jj} - 2c_{ij} \geqslant 0 \qquad \forall i \neq j \in N.$$

By using the equations

$$c_{ii} = -\sum_{j \neq i} c_{ij} \qquad \forall i \in N,$$

we can eliminate the c_{ii} from the inequalities, and obtain $\frac{1}{2}n(n-1)$ homogeneous inequalities with $\frac{1}{2}n(n-1)$ unknowns. In our standard example they are

$$4c_{12} + c_{13} + c_{23} \geqslant 0$$
$$c_{12} + 4c_{13} + c_{23} \geqslant 0$$
$$c_{12} + c_{13} + 4c_{23} \geqslant 0.$$

The matrix

$$\begin{matrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{matrix}$$

is nonsingular, in fact it is even PD with spectrum (6 3 3). The cone is thus pointed, and we can solve for the edges.

$$4c_{12} + c_{13} + c_{23} = 0$$
$$c_{12} + 4c_{13} + c_{23} = 0.$$

Subtract

$$3c_{12} - 3c_{13} = 0,$$

or

$$c_{12} = c_{13}.$$

Substitute

$$5c_{12} + c_{23} = 0,$$

or

$$c_{23} = -5c_{12}.$$

The first edge has the form $(c_{12}\ c_{13}\ c_{23}) = (\ a\ \ a\ -5a)$. The final inequality demands that $a + a - 20a \geqslant 0$, and consequently $a \geqslant 0$. The edge is

$$\begin{matrix} -2 & 1 & 1 \\ 1 & 4 & -5 \\ 1 & -5 & 4 \end{matrix}$$

And by symmetry the other edges are

$$\begin{matrix} 4 & 1 & -5 \\ 1 & -2 & 1 \\ -5 & 1 & 4 \end{matrix} \qquad \begin{matrix} 4 & -5 & 1 \\ -5 & 4 & 1 \\ 1 & 1 & -2 \end{matrix}$$

All three values of r are equal to -.20, the spectra are (9 0 -3). Next we may apply the n necessary conditions $c_{ii} \geqslant 0$. The possibility is clear: for n = 1(1)20 for example we may make a standard library of edges. Again the decision must be made how many necessary conditions these edges must obey. Adding another condition means higher intercorrelations (a smaller cone, a closer approximation to the cone of PSD matrices), but it also means more edges. If something like this was done, the program that takes care of the first stage would read in the ready-made edges, apply the inequalities $Tr(B_k C) \geqslant 0$ one by one, and deliver the new edges of Q_C for the second stage.

Some remarks must be made on the treatment of ties ($\delta_{ij} =_o \delta_{kl}$). If we require semi-strong order-isometricity, then we must require $d^2_{ij} = d^2_{kl}$, and this defines an equation in C. Weak order-isometricity means that we require nothing about the relation between d_{ij} and d_{kl}, there simply is no inequality in this case. Both approaches can, of course, readily be incorporated in the first stage. The semi-strong approach makes the cones Q_s (and P) smaller, and the values of r_s larger. For the weak approach the reverse is true.

Consider the partial order

$$\delta_{12} \geqslant_o \delta_{23}$$
$$\delta_{13} \geqslant_o \delta_{23}$$

which contains less information then the weak order $\delta_{12} \geqslant_o \delta_{13} \geqslant_o \delta_{23}$ we investigated previously. If we use the corners we have already computed (i.e. the ones on this page) and apply the two inequalities from the data then we obtain edges

$$\begin{matrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{matrix} \qquad \begin{matrix} 4 & 1 & -5 \\ 1 & -2 & 1 \\ -5 & 1 & 4 \end{matrix} \qquad \begin{matrix} 4 & -5 & 1 \\ -5 & 4 & 1 \\ 1 & 1 & -2 \end{matrix}$$

(1) (2) (3)

The value of r_{min} is still $-.20$. The edge (3) has spectrum (9 0 -3), and with -3 goes the vector (-1 -1 2). New inequality $Tr(CD_1) \geqslant 0$, with D_1 =

$$\begin{matrix} 1 & 1 & -2 \\ 1 & 1 & -2 \\ -2 & -2 & 4 \end{matrix}$$

The values of d_i are (18 36 -18), which gives us as new edges

$$\begin{matrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{matrix} \qquad \begin{matrix} 4 & 1 & -5 \\ 1 & -2 & 1 \\ -5 & 1 & 4 \end{matrix} \qquad \begin{matrix} 4 & -3 & -1 \\ -3 & 2 & 1 \\ -1 & 1 & 0 \end{matrix} \qquad \begin{matrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{matrix}$$

(1) (2) (3) (4)

and r_{min} has gone up to $r_{24} = 0$. Cut off (2) with $Tr(CD_2) \geqslant 0$, with D_2 =

$$\begin{matrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{matrix}$$

Values for d_i (18 -18 18 9), new edges

$$\begin{matrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{matrix}$$

(1)

$$\begin{matrix} 4 & -3 & -1 \\ -3 & 2 & 1 \\ -1 & 1 & 0 \end{matrix} \qquad \begin{matrix} 4 & -1 & -3 \\ -1 & 0 & 1 \\ -3 & 1 & 2 \end{matrix}$$

(2) (3)

$$\begin{matrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{matrix} \qquad \begin{matrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{matrix}$$

(4) (5)

$$\begin{matrix} 6 & -1 & -5 \\ -1 & 0 & 1 \\ -5 & 1 & 4 \end{matrix}$$

(6)

R:	(1)	1.000	.655	.655	.707	.707	.690
	(2)		1.000	.852	.926	.463	.632
	(3)			1.000	.463	.926	.700
	(4)				1.000	.250	.620
	(5)					1.000	.900
	(6)						1.000

Spectra:	(1)	3	3	0
	(2)	$3 + 2\sqrt{3}$	0	$3 - 2\sqrt{3}$
	(3)	$3 + 2\sqrt{3}$	0	$3 - 2\sqrt{3}$
	(4)	2	0	0
	(5)	2	0	0
	(6)	$5 + 2\sqrt{2}$	$5 - 2\sqrt{2}$	0

This example illustrates some important points. In the first place investigating the partial order makes clear that, in any order-isometric representation, interchanging the points x_2 and x_3 is permitted. Therefore edges like (2)-(3), and (4)-(5) occur in pairs in which the second and third row and column are permuted. The second important point is that the minimum value of r in this cycle is obtained for edges (4) and (5), and that both these edges are PSD. It follows that the value of .250 is the minimum value of r for all pairs of matrices $C_1, C_2 \in P$ (convergence in just two steps). The representation corresponding to (4) is identical to the one in figure 1a, the representation corresponding to (5) is the same but with x_2 and x_3 interchanged. One-dimensional procedures would give us the same configurations with minimum PM-correlation of .500. Finally, while $r_{min} \geq .65$ in the weakly ordered case, deleting one inequality makes r_{min} equal to .250. This has a moral. It should make one a little bit pessimistic about the uniqueness of scaling solutions with categorical data (such as those collected with sociogram-type techniques and multiple choice tests). Not too pessimistic however, because we did not yet study the effect of an increasing number of points on r_{min}. It is not true that $r_{min} \rightarrow 1$ if $n \rightarrow \infty$ (cf section 3.9). The degenerate solutions turn out to be the big nuisance once again. There are two ways to circumvent this. We may cling to strong order-isometricity, but it is numerically very difficult to work with open convex cones. The numerically more feasible way is to have an objective criterion for selecting a matrix in the cone which is, in a mathematically well-defined sense, the best one. This approach is discussed in chapter 4.

3.7 A somewhat different approach

Suppose that all $\binom{n}{2}$ upper-diagonal dissimilarities are weakly ordered. This gives us $\binom{n}{2} - 1$ inequalities of the form

$$d^2_{ij} - d^2_{kl} \geqslant 0,$$

and we use an extra one (for the smallest δ, say δ_{pq})

$$d^2_{pq} \geqslant 0.$$

Consider the $\binom{n}{2}$-dimensional space in which each d^2_{ij} defines an axis. The inequalities define a pointed cone with $\binom{n}{2}$ edges

$$\begin{array}{l} (0\ 0\ \ldots\ 0\ 1) \\ (0\ \ldots\ 0\ 1\ 1) \\ \ldots\ldots\ldots \\ (0\ 1\ \ldots\ 1\ 1) \\ (1\ 1\ \ldots\ 1\ 1). \end{array}$$

Now consider $\binom{n}{2}$ nonnegative quantities x_p, and let r_{ij} denote the rank number of δ_{ij}. Then a vector $\{d^2_{ij}\}$ lies in the cone iff there is another vector x such that

$$d^2_{ij} = \sum_{p=1}^{r_{ij}} x_p.$$

For all vectors x we can define a matrix V with

$$v_{ij} = \sum_{p=1}^{r_{ij}} x_p, \qquad \forall i \neq j,$$

$$v_{ii} = 0, \qquad \forall i,$$

and a matrix W with

$$w_{ij} = n \sum_{k=1}^{n} v_{ik} + n \sum_{k=1}^{n} v_{kj} - \sum_{k=1}^{n} \sum_{l=1}^{n} v_{kl} - n^2 v_{ij}.$$

It is well known that V can be interpreted as a matrix of squared Euclidean distances iff W is PSD (it is always true that W is SDC). If we substitute the expression for v_{ij} in the definition of W we obtain a matrix whose elements are linear functions of the x_p. We require that V is monotone, or equivalently, $x_p \geqslant 0$. This defines the edges

$$\begin{array}{c} (1\ 0\ 0\ .\ .\ .\ 0) \\ (0\ 1\ 0\ .\ .\ .\ 0) \\ .\ .\ .\ .\ .\ .\ . \\ (0\ .\ .\ .\ 0\ 0\ 1), \end{array}$$

and by substituting these edges into the expressions for w_{ij} we obtain the edges of Q_0. In our familiar example the matrix V looks like

$$\begin{array}{ccc} 0 & x_1+x_2+x_3 & x_1+x_2 \\ x_1+x_2+x_3 & 0 & x_1 \\ x_1+x_2 & x_1 & 0 \end{array}$$

which gives for W

$$\begin{array}{ccc} 6x_1+8x_2+4x_3 & -3x_1-4x_2-5x_3 & -3x_1-4x_2+x_3 \\ -3x_1-4x_2-5x_3 & 6x_1+2x_2+4x_3 & -3x_1+2x_2+x_3 \\ -3x_1-4x_2+x_3 & -3x_1+2x_2+x_3 & 6x_1+2x_2-2x_3. \end{array}$$

Substitute the edges

$$\begin{array}{ccc} x_1 & x_2 & x_3 \\ (\ 1 & 0 & 0) \\ (\ 0 & 1 & 0) \\ (\ 0 & 0 & 1), \end{array}$$

which gives

$$\begin{array}{ccc} \begin{array}{ccc} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{array} & \begin{array}{ccc} 8 & -4 & -4 \\ -4 & 2 & 2 \\ -4 & 2 & 2 \end{array} & \begin{array}{ccc} 4 & -5 & 1 \\ -5 & 4 & 1 \\ 1 & 1 & -2 \end{array} \\ (1) & (2) & (3) \end{array}$$

The extra requirement $w_{33} \geqslant 0$ cuts off (3) and replaces it with two others

$$\begin{array}{cc} \begin{array}{ccc} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{array} & \begin{array}{ccc} 4 & -3 & -1 \\ -3 & 2 & 1 \\ -1 & 1 & 0 \end{array} \\ (3') & (3'') \end{array}$$

The rationale seems different, but is essentially identical to the one we used previously. We use

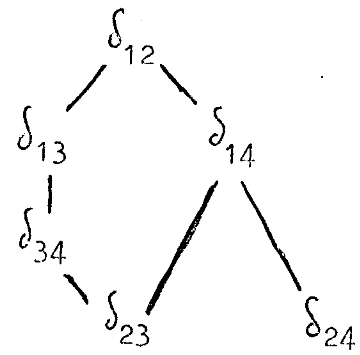
$$\underset{C\ \ \ \mathrm{SDC}}{\mathrm{Tr}(B_k C) \geqslant 0}$$

and an additional inequality that makes the smallest $d_{pq}^2 = c_{pp} + c_{qq} - 2c_{p}$ 0. Observe that for these edges it is automatically true that

$$c_{ii} + c_{jj} - 2c_{ij} \geqslant 0$$

for all i,j (which are $\binom{n}{2}$ necessary conditions).

We have transformed the problem back to a different space in which each of the d_{ij}^2 with $i < j$ defines an axis. In this new space the edges turn out to be much easier to compute. The inequalities $d_{ij}^2 \geqslant d_{kl}^2$ reduce to $x_p \geqslant 0$, and the equations $d_{ij}^2 = d_{kl}^2$ to $x_p = 0$. For partial orders the situation is somewhat more complicated. If we analyze the partial order



we end up with about 25 x_p-variables (and consequently about 25 edges) in stead of six. Nevertheless, even in this case, the method is still much more simple than the ones in the previous section.

3.8 The role of nI-E

From the examples we have investigated so far it is clear that the matrix nI − E plays a peculiar role. In the first place it is always a solution (in most cases a weak (n-1)-solution, it is strong or semi-strong iff $\geqslant_o$ is an equivalence relation that connects the nondiagonal δ_{ij}), in the second place it is always an edge of the cone P. Using hI − E we can give upper bounds for the value of r_p.

D5

Theorem 3.7: If our system of inequalities has a weak k-solution then

$$r_p \leq \sqrt{\frac{k}{n-1}} .$$

Proof: Suppose X is a weak k-solution. Without loss of generality we assume that $X'X = \Lambda$ is diagonal, and that the columns of X are centered. Write C for nI - E. Then

$$\mathrm{Tr}(CXX') = \sum_{i=1}^{k} x'_{.i} Cx_{.i} = n \sum_{i=1}^{k} \lambda_{ii},$$

and

$$\mathrm{Tr}(XX'XX') = \mathrm{Tr}(\Lambda^2) = \sum_{i=1}^{k} \lambda_{ii}^2 .$$

Of course

$$\mathrm{Tr}(C^2) = n(n-1)^2 + n(n-1) = n^2(n-1),$$

and thus

$$r_p \leq r(C,XX') = \frac{\sum \lambda_{ii}}{\sqrt{(n-1)\sum \lambda_{ii}^2}} \leq \frac{\sqrt{k \sum \lambda_{ii}^2}}{\sqrt{(n-1)\sum \lambda_{ii}^2}} = \sqrt{\frac{k}{n-1}} .$$

Q.E.D.

<u>Corollary 3.5</u>: If the system has a weak 1-solution then

$$r_p \leq r(Cxx') = \sqrt{\frac{1}{n-1}} .$$

The next thing we investigate is what exactly the influence is of adding nI - E to a particular solution.

<u>Lemma 3.4</u>: If C is a SDC matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}, 0$ and $\tilde{C} = C + \alpha(nI - E)$, then C and $\tilde{C}$ have the same eigenvectors. $\tilde{C}$ has the eigenvalues zero and $\lambda_i + \alpha n$ $(i=1,\ldots,n-1)$.

Proof: Both C and $\tilde{C}$ are doubly centered, therefore $Ce = \tilde{C}e = 0$, and all ot eigenvectors are centered. If $Cx = \lambda_i x$ with $x \neq e$ then $\tilde{C}x = Cx + \alpha(nI -$ $= \lambda_i x + \alpha nx = (\lambda_i + \alpha n)x$. If $\tilde{C}x = \gamma_i x$ with $x \neq e$ then $Cx = \tilde{C}x -$ $\alpha(nI - E)x = \gamma_i x - \alpha nx = (\gamma_i - \alpha n)x$. Q.E.D.

In terms of the representation: if $C = K\Lambda K'$, or $X = K\Lambda^{\frac{1}{2}}$, then $C =$ $K(\Lambda + \alpha nI)K'$, and $\tilde{X} = K(\Lambda + \alpha nI)^{\frac{1}{2}}$. Interms of the distances: if $d_{ij}^2 =$ $\sum_s \lambda_s(k_{is} - k_{js})^2$, then $\tilde{d}_{ij}^2 = \sum_s (\lambda_s + \alpha n)(k_{is} - k_{js})^2 = \sum_s \lambda_s(k_{is} - k_{j}$

$+ \quad n \sum_s (k_{is} - k_{js})^2 = d_{ij}^2 + 2\alpha n$. This can also be proved by using $\tilde{d}_{ij}^2 = \tilde{c}_{ii} + \tilde{c}_{jj} - 2\tilde{c}_{ij} = c_{ii} + c_{jj} - 2c_{ij} + \alpha\,(2(n-1) + 2) = d_{ij}^2 + 2\alpha n$.

Theorem 3.8: If C is a solution then $\tilde{C} = C + \alpha\,(nI - E)$ is a solution. In this new solution

i) all squared distances increase by the same positive constant

ii) the principal axes of C are differentially stretched.

For each solution C we may now solve the problem

$$\alpha \quad \max!$$

$$C - \alpha\,(nI - E) \qquad \text{PSD.}$$

From lemma 3.4 it follows that if C is an (n-1)-solution with smallest positive eigenvalue $\tilde{\lambda}$, then the solution $\tilde{\alpha} = \tilde{\lambda}/n$. If C is of rank less than n - 1, then $\tilde{\alpha} = 0$. Finally, if C is $\beta\,(nI - E)$, then $\tilde{\alpha} = \beta$. The matrix

$$\tilde{C} = C - \tilde{\alpha}\,(nI - E)$$

is called the α-canonical form of C. Evidently $\mathrm{rank}(\tilde{C}) \leq n - 2$, and the α-canonical form of nI - E is 0. Because nI - E is always a solution, we have interpreted it as just another type of indeterminacy that has to be eliminated. And now we may say that we are only interested in scaled SDC solutions in α-canonical form.

Theorem 3.9: The edges of the cone P (except nI - E) are always in α-canonical form.

Proof: If not then there is a solution $\tilde{C}$ and a positive number $\tilde{\alpha}$ such that

$$\tilde{C} + \tilde{\alpha}\,(nI - E) = C$$

which means that C is not an edge. Q.E.D.

Corollary 3.6: All edges of P (except nI - E) have rank $\leq n - 2$.

If we apply the procedure of the previous section to the example $\delta_{12} \geq \delta_{1}$ $\delta_{14} \geq \delta_{23} \geq \delta_{24} \geq \delta_{34}$ we find the edges

$$\begin{array}{rrrr} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{array} \qquad \begin{array}{rrrr} 7 & -1 & -3 & -3 \\ -1 & 7 & -3 & -3 \\ -3 & -3 & 3 & 3 \\ -3 & -3 & 3 & 3 \end{array} \qquad \begin{array}{rrrr} 4 & -1 & -1 & -2 \\ -1 & 2 & -2 & 1 \\ -1 & -2 & 2 & 1 \\ -2 & 1 & 1 & 0 \end{array}$$

$$(1) \qquad\qquad\qquad (2) \qquad\qquad\qquad (3)$$

$$\begin{array}{rrrr} 9 & -3 & -3 & -3 \\ -3 & 1 & 1 & 1 \\ -3 & 1 & 1 & 1 \\ -3 & 1 & 1 & 1 \end{array} \qquad \begin{array}{rrrr} 3 & -2 & -2 & 1 \\ -2 & 1 & 1 & 0 \\ -2 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{array} \qquad \begin{array}{rrrr} 3 & -5 & 1 & 1 \\ -5 & 3 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{array}$$

$$(4) \qquad\qquad\qquad (5) \qquad\qquad\qquad (6)$$

Edges (1), (2), (4) are PSD. Edge (4) is of rank one, which means (by corollary 3.5) $r_{14} = \sqrt{\frac{1}{3}}$. The configuration is given in figure 2a. Edge (2 is of rank two, which means (theorem 3.7) $r_{12} \leq \sqrt{\frac{2}{3}}$. In fact the eigenvalu are 12 and 8 and the proof of the theorem shows that $r_{12} = 20/\sqrt{624} \approx .8$. The configuration is given in figure 2b. Both solution are, of course, mor or less degenerate, and both solutions are in α -canonical form.

Another example starts from the distances

$$\begin{array}{cccc} 0 & a\sqrt{2} & 2a & a\sqrt{2} \\ & 0 & a\sqrt{2} & 2a \\ & & 0 & a\sqrt{2} \\ & & & 0 \end{array}$$

obtained from the configuration in figure 2c. Requiring semi-strong order-isometricity gives us the two edges

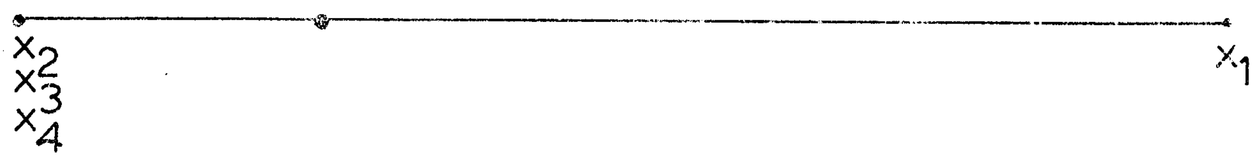
$$\begin{array}{rrrr} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{array} \qquad \begin{array}{rrrr} 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \\ -3 & 1 & -3 & 1 \\ 1 & -3 & 1 & 1 \end{array}$$

$$(1) \qquad\qquad\qquad (2)$$

Edge (2) has the following proper vectors and values

$$\begin{array}{rrrr} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 1 \\ \hline 4 & 4 & 0 & -4 \end{array}$$

fig 2a



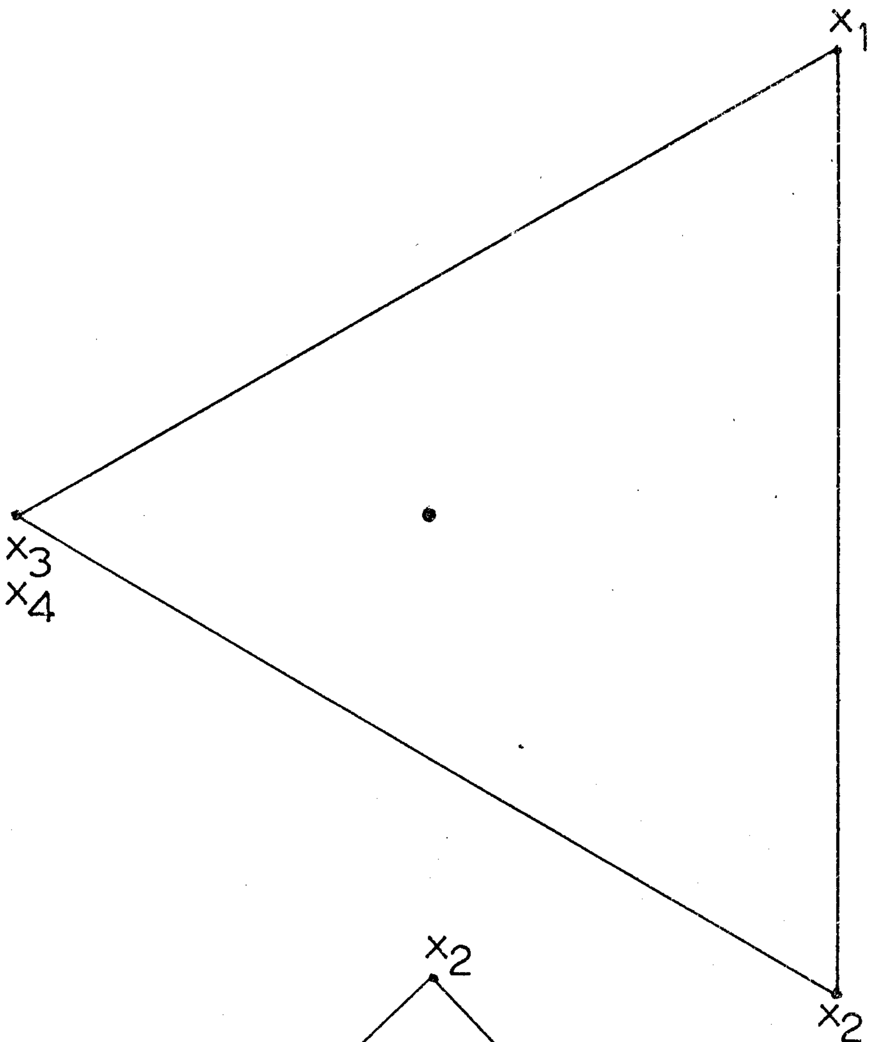


fig 2b

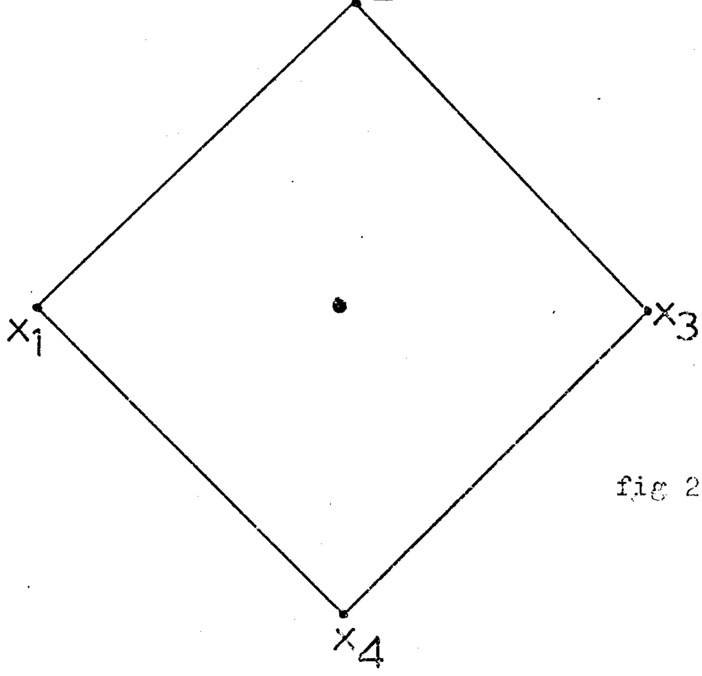


fig 2c

Observe that the first two vectors already reproduce the configuration exactly. Nevertheless the edge is not PSD, we cut it off, and obtain the new edge

$$\begin{matrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{matrix}$$

(2')

which is PSD (and in α-canonical form) with spectrum (2 2 0 0), and with eigenvectors equal to those of (2). We have applied lemma 3.4 with $\alpha = 1$. In this case the two eigenvalues are equal, which is the necessary and sufficient condition for equality in theorem 3.7, and thus $r_p = \sqrt{\frac{2}{3}}$. If we consider nI - E as something redundant, we can say that the solution is essentially unique. In this particular case the matrix $\tilde{C} = C + \alpha\,(nI - E)$ has eigenvectors equal to those of (2) and eigenvalues ($2 + 4\alpha$ $\;2 + 4\alpha$ $\;4\alpha$ $\;0$), which means that the first two dimensions still reproduce the configuration perfectly, but that there also exists a third nuisance dimension with positive eigenvalue. From the point of view of minimum dimensionality (or: if we are looking for the maximally parsimonious solution) it seems quite natural to restrict our attention to α-canonical solutions.

The fact that nI - E is among the solutions can, incidentally, also be used to apply the supporting-hyperplane version of the cutting methods to our problem (Veinott 1967; Zangwill 1969, ch 14). If $\bar{C}$ is the solution of a linearized subproblem with minimum eigenvalue $\bar{\lambda} < 0$, we find a point $\hat{C}$ on the boundary by

$$\tilde{C} = \bar{C} - \frac{\bar{\lambda}}{n}\,(nI - E).$$

$\hat{C}$ is PSD, because by lemma 3.4 $\lambda_{min}(\hat{C}) = \lambda_{min}(\bar{C}) - \bar{\lambda} = \bar{\lambda} - \bar{\lambda} = 0$. The next supporting hyperplane is

$$\mathrm{Tr}(\overline{yy}'(C - \tilde{C})) = \mathrm{Tr}(C\bar{y}\,\bar{y}') - \bar{y}'\tilde{C}\bar{y} = \mathrm{Tr}(C\bar{y}\,\bar{y}') = 0,$$

which is the same as the cutting plane constructed in the concave cutting plane method. In this particular case the two methods are equivalent.

The results we obtained so far make it possible to prove a considerably more precise representation theorem. For the moment we assume that ϕ is a POG-distance, and we even start with embedding it in a WOG-distance. In this way we obtain a weak order over all upper-diagonal dissimilarities δ_{ij} with $j > i$. The elements of the matrix W in section 3.7 are linear functions of the $\binom{n}{2}$ nonnegative variables x_p, and

$$Q_0 = \left\{ C \;\middle|\; C = \sum_{p=1}^{\binom{n}{2}} x_p T_p \;,\; x_p \geq 0 \right\},$$

where the T_p are $\binom{n}{2}$ SDC-matrices. In our example in section 3.7 the expression for W gives

$$\underset{T_1}{\begin{matrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{matrix}} \qquad \underset{T_2}{\begin{matrix} 8 & -4 & -4 \\ -4 & 2 & 2 \\ -4 & 2 & 2 \end{matrix}} \qquad \underset{T_3}{\begin{matrix} 4 & -5 & 1 \\ -5 & 4 & 1 \\ 1 & 1 & -2 \end{matrix}}$$

If we require $d^2_{ij} = d^2_{kl}$ we simply set one of the x_p equal to zero and the corresponding matrix T_p is not included in the summation (requiring $x_1 = 0$ is equivalent to requiring $d^2_{pq} = 0$ if δ_{pq} is the smallest of the δ_{ij}). We eliminate the T_p corresponding with values of x_p that must vanish. This gives

$$Q_0 = \left\{ C \;\middle|\; C = \sum_{p=1}^{k} x_p T_p \;,\; x_p \geq 0 \right\},$$

with $k \leq \binom{n}{2}$. All elements of the cone Q_0 are monotone in the semi-strong sense. If we demand $x_p > 0$ for $p = 1,\ldots,k$ then they are even monotone in the strong sense. Let

$$Q_0^i = \left\{ C \;\middle|\; C = \sum_{p=1}^{k} x_p T_p \;,\; x_p > 0 \right\}.$$

<u>Lemma 3.5</u>: If one of the T_p is not PSD, then Q_0^i has an element which is not PSD.

Proof: Suppose $\lambda_{min}(T_k) < 0$. Consider any point $\overline{C} = \sum_{p=1}^{k-1} x_p T_p$ with $x_p > 0$.

Then the point

$$\tilde{C} = \mu \overline{C} + (1 - \mu) T_k$$

belongs to Q_0^i for $0 < \mu < 1$. Along this line $\lambda_{min}(\tilde{C})$ is a continuous function of μ. No matter whether $\lambda_{min}(\overline{C}) < 0$ or $\lambda_{min}(\overline{C}) = 0$, it follows by continuity that there is a $0 < \mu < 1$ such that $\lambda_{min}(\tilde{C}) < 0$. Q.E.D.

Lemma 3.6: If $C \in Q_0^i$ and $\lambda_{min}(C) = \overline{\lambda} < 0$, then $\overline{C} = C - \frac{\overline{\lambda}}{n}(nI - E) \in Q_0^i$ and $\text{rank}(\overline{C}) \leqslant n - 2$.

Proof: The first part is obvious, the second part follows from lemma 3.4. Q.E.D.

Our new representation theorem uses the following definition: a POG-distance is nondegenerate iff $\geqslant_0$ is not an equivalence relation that connects the δ_{ij} with $j > i$.

Theorem 3.10: The Euclidean NMS-problem has a strong (n-2)-solution $\overline{C}$ iff ϕ is a nondegenerate POG-distance.

Proof: Take a positive combination C of the k matrices T_p found by requiring semi-strong order-isometricity and positivity of the d_{ij}^2. Take it in such a way that $C \neq nI - E$ (if ϕ is nondegenerate this can be done). If C is PSD of rank $\leqslant n - 2$, then we are ready: $\overline{C} = C$. If C is PSD and $\text{rank}(C) = n - 1$, then $\overline{C}$ is the α-canonical form of C. If C is not PSD, then we find $\overline{C}$ by applying lemma 3.6. Thus the condition is sufficient. Necessity follows easily from definition 0.1. Q.E.D.

This representation theorem seems somewhat stronger that Guttman's SSA-I theorem. Guttman's proof is not published as yet, and it is quite possible that he uses quite different methods. The two existence theorems mentioned by Guttman in his basic NMS-paper (Guttman 1968 p 476-477) are corollaries of our theorem 3.10. Our representation theorem 2.4 is not exactly a corollary of 3.10. A third form of the representation theorem can be found in a companion paper (De Leeuw: Metric methods in Euclidean NMS, in preparation).

Of course these theorems can hardly be compared with the representation theorem of Beals, Krantz, and Tversky (cf also Beals & Krantz 1968, Krantz 1968, Krantz & Tversky 1969). Their theorem is meant for the infinite, idealized case, and their sufficient conditions have the essentially different status of underlying scientific assumptions, which can (and must) be tested by experiments. Our necessary and sufficient condition is merely a mathematical nicety without any deeper scientific meaning. Theorem 3.10 states, crudely, that the Euclidean MDS-model is more or less tautological in the finite case, because its requirements can almost always be satisfied for $p = n - 2$. Although the fact that $nI - E$ is always present as a solution may be called a nuisance for computational work, it is very important from a theoretical point of view. It has given us some interesting theorems. As a final corollary of theorems 3.7 and 3.10 we have

<u>Corollary 3.7</u>: If ϕ is a nondegenerate POG-distance then

$$r_P \leqslant \sqrt{\frac{n-2}{n-1}}.$$

We investigate two examples in which our POG-distance is almost degenerate. The first one is $\delta_{12} =_o \delta_{13} =_o \delta_{14} =_o \delta_{23} =_o \delta_{24} \lesssim \delta_{34}$. We find

$$\begin{matrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{matrix} \qquad \begin{matrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 3 & -5 \\ 1 & 1 & -5 & 3 \end{matrix}$$

as the two edges (T-matrices) of Q_0. The (strictly monotone) combination $\alpha T_1 + \beta T_2$ has roots and vectors

eigenvector:	0	0	1	-1	eigenvalue:	$4\alpha + 8\beta$
	1	-1	0	0		4α
	-1	-1	1	1		$4\alpha - 4\beta$
	1	1	1	1		0

Taking $\alpha = \beta = \frac{1}{4}$ gives us the strong 2-solution in figure 3a. If we take $\alpha > \beta > 0$, then we also obtain solutions, but they are not in α-canonical form and of rank 3. The second example is $\delta_{12} =_o \delta_{13} =_o \delta_{14} =_o \delta_{23} =_o \delta_{24}$ (there is no information about δ_{34}).

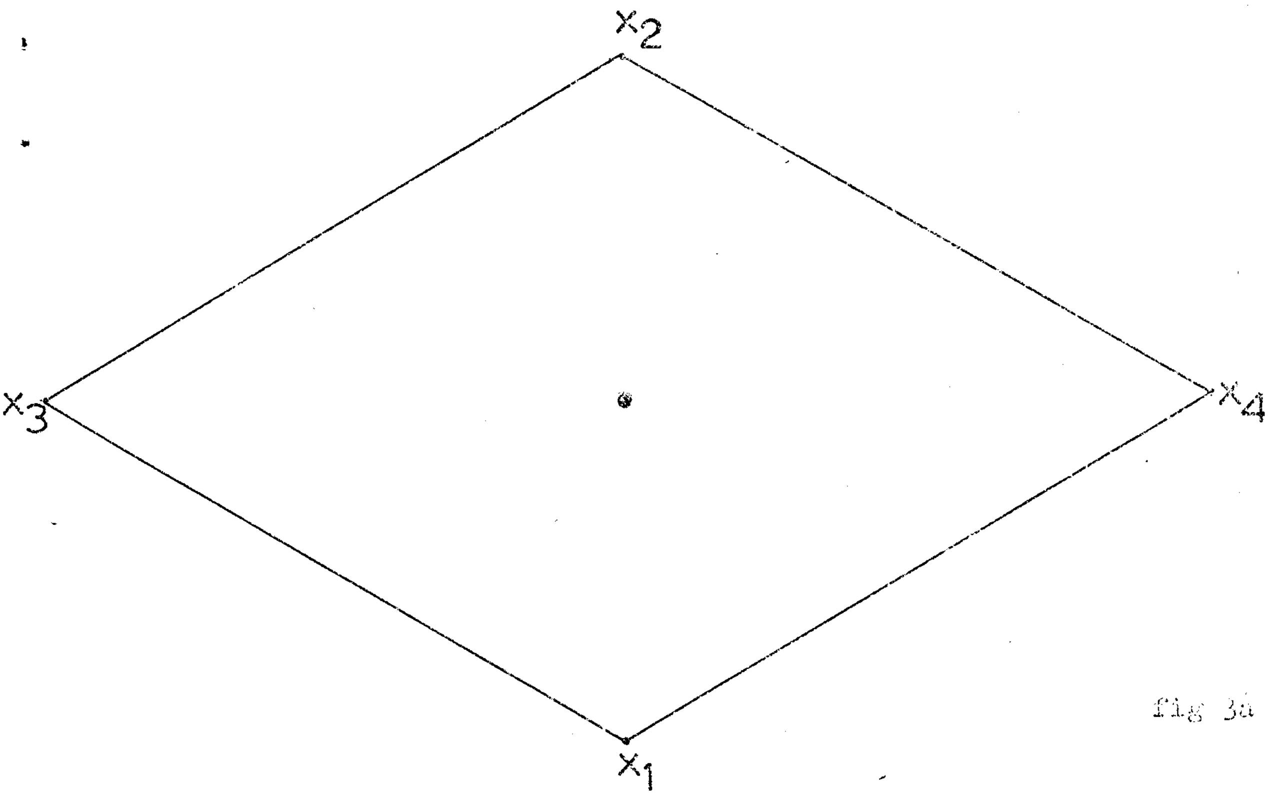


fig 3a

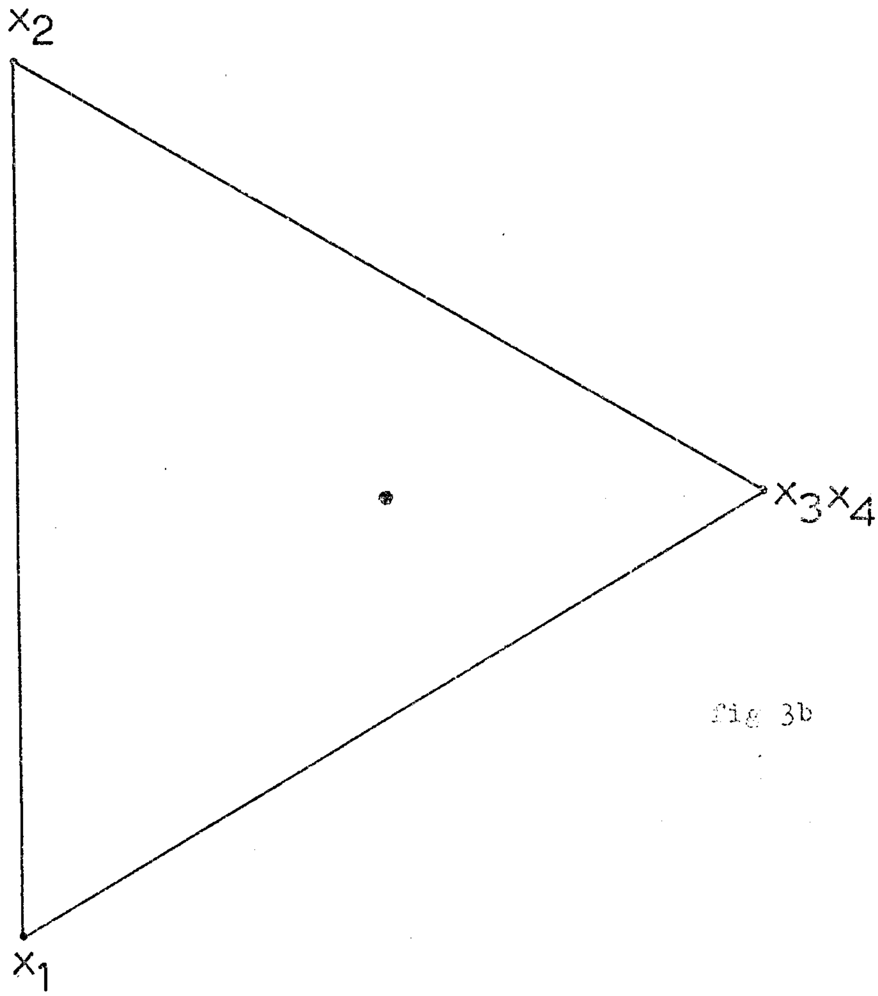


fig 3b

The two edges are

$$\begin{matrix} 7 & -1 & -3 & -3 \\ -1 & 7 & -3 & -3 \\ -3 & -3 & 3 & 3 \\ -3 & -3 & 3 & 3 \end{matrix} \qquad \begin{matrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 3 & -5 \\ 1 & 1 & -5 & 3 \end{matrix}$$

(observe that in this case $\frac{1}{2}T_1 + \frac{1}{2}T_2 = nI - E$, which means that $nI - E$ is not an edge !). Again we investigate the roots and vectors of $\alpha T_1 + \beta T_2$.

eigenvector:	1 −1 0 0	eigenvalue: 8α
	0 0 1 −1	8β
	−1 −1 1 1	$12\alpha - 4\beta$
	1 1 1 1	0

In this case both T_1 (configuration in figure 3b) and $\frac{1}{4}T_1 + \frac{3}{4}T_2$ (configuration identical to the one in figure 3a) are strong 2-solutions. The value of r_p is approximately .557. Observe that in the first configuration d_{34} is the largest of the distances (fig 3a), in the second one it is the smallest (fig 3b). Both solutions are, of course, in α-canonical form.

3.9 Further computational suggestions

The method of section 3.7, used there to compute the edges of Q_0 more efficiently, can also be used to reformulate our problem. The NMS-problem is to find values of x_p such that

$$\sum_p^k x_p T_p \qquad \text{PSD},$$

$$x_p \geqslant 0,$$

where the T_p are the k different SDC matrices defined in section 3.8. Indeterminacy due to uniform stretching and shrinking can be eliminated by requiring

$$\mathrm{Tr}\left(\sum_p^k x_p T_p\right) = \sum_p^k x_p \mathrm{Tr}(T_p) = n(n-1).$$

Solving for one of the x_p (without loss of generality x_k) gives

$$\sum_{p=1}^{k-1} x_p S_p + Q \qquad \text{PSD},$$

$$c'x \leqslant n(n-1),$$

$$x_p \geqslant 0.$$

The problem of linearizing the constraint set is of the form

$$F(x) = \lambda_{min}(\sum x_p S_p + Q) \quad \text{min!}$$

$$c'x \leqslant n(n-1),$$

$$\sum x_p S_p + Q \quad \text{PSD},$$

$$x_p \geqslant 0,$$

and the linearized subproblems become

$$F(x) \quad \text{min!}$$

$$c'x \leqslant n(n-1),$$

$$Ax \leqslant b,$$

$$x \geqslant 0.$$

The linear restrictions $Ax \leqslant b$ are generated by the cutting constraints

$$\sum x_p y_s' S_p y_s + y_s' Q y_s \geqslant 0,$$

or

$$\sum x_p(-y_s' S_p y_s) \leqslant y_s' Q y_s.$$

The form of these linearized subproblems makes it possible to use the ordinary simplex tableau. Furthermore we can use the terms basic solution, basic feasible solution, pivot step, etc. as in the simplex method (for the moment we forget about the relative-cost vector and the value of the objective function). Obviously $x = 0$ defines a basic solution (in general not feasible) with tableau

$$z = \begin{bmatrix} n(n-1) \\ \cdots\cdots \\ b \end{bmatrix} \begin{array}{c} [\; x = 0 \;] \\ \begin{bmatrix} c' \\ \cdots\cdots\cdots\cdots \\ A \end{bmatrix} \end{array}$$

In the case that γ_0 is connected and x_1 corresponds with the smallest δ_{ij} it is easy to see that the (basic feasible) solution corresponding with $x_1 > 0$, $x_2 = x_3 = \ldots = x_{k-1} = 0$ is $nI - E$. If we perform one pivot step which takes z_1 out of the basis and x_1 into the basis, then we obtain a basic feasible solution. If A contains l columns, then the number of basic solutions is not larger than

$$\binom{k-1+l+1}{k-1} = \binom{k+l}{k-1} = \binom{l+1}{k-1}.$$

Quite a number of these vertices will not be feasible.

This suggests a first algorithm. Construct the tableau, investigate all basic solutions (vertices); if they are feasible then compute the corresponding matrix C and its λ_{min}, retain only the smallest λ_{min} with associated eigenvector. This procedure has already a few obvious advantages over the ones we outlined previously. In the first place we do not need much storage, in the second place we are sure that the procedure does not generate feasible matrices which are not vertices. The amount of computation, however, may still be enormous. Unfortunately alternative methods which are sure to find all vertices (and nothing but vertices) do not seem to be available. The method that comes closest is the one due to Balinski (1961), which is, in a sense, the dual of the double description method used by Motzkin et al, Uzawa, Lindman, and by us in section 3.2 and further. Balinski's method is quite simple. If the polyhedron in n-space is the solution set of $Ax \leq b$, then start by picking out one of the $a_{i.}$, and find all vertices on the hyperplane $a_{i.}x = b_i$. We have reduced the problem to one in (n-1)-space. This problem can be reduced to one in (n-2)-space, and so on, until we reach two-space in which the problem is particulary simple and can be solved easily without further reduction. If we have found all the vertices on the hyperplane we delete the constraint $a_{i.}x \leq b_i$, and we start all over again on this new polyhedron in n-space. The method does, of course, find all vertices of the original polyhedron, but it also finds all vertices of the polyhedra that are formed in the course of the computation by deleting hyperplanes. As an example of how the method works we study the polyhedron in figure 4a. First study the hyperplane (A), and find (1) and (2). Deleting the constraint corresponding with (A) means forming the new polyhedron in 4b. Take (B) and find (5) and (6), observe that (6) is not a vertex of our original polyhedron. Delete (B), pick (C) in the (unbounded) polyhedron 4c, find (4). A final step gives 4d and (3). Balinski reports quite satisfactory results with this method, even in a comparatively large problem (65 nonnega-

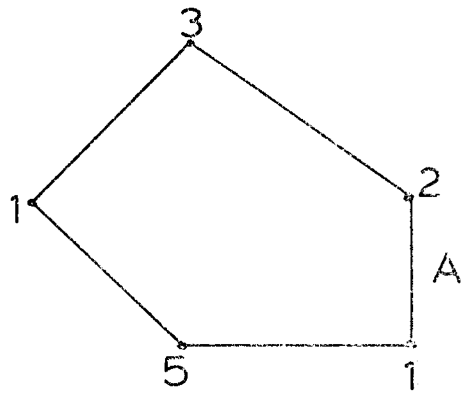


fig 4a

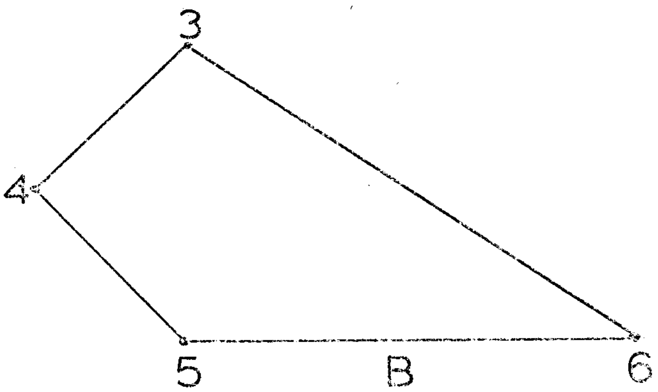


fig 4b

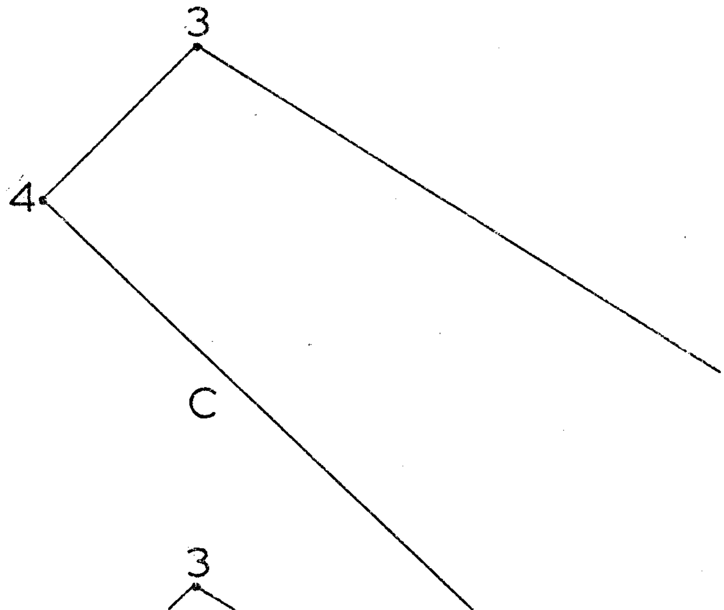


fig 4c

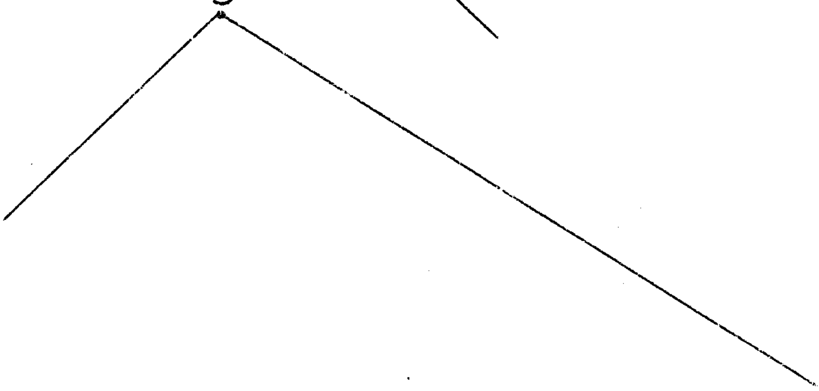


fig 4d

tive variables, 35 restrictions, 31 vertices). In our problem the situation is much more complicated, because none of the linear restrictions is redundant.

An alternative procedure is to linearise our objective function. Suppose we have found a basic feasible point $\bar{x}$ (with associated $\bar{C}$, $\bar{\lambda}$, and $\bar{y}$). Linearize λ_{min} in that point

$$\bar{F}(x) = \sum x_p \bar{y}' S_p \bar{y} + \bar{y}' Q \bar{y}.$$

We know that $\bar{F}(\bar{x}) = \bar{\lambda}$. Now solve the 'completely linearised' problem

$$\bar{F}(x) \quad \text{min!}$$

$$c'x \leqslant n(n-1),$$

$$Ax \leqslant b,$$

$$x \geqslant 0,$$

by some variant of the simplex method.

<u>Theorem 3.11</u>: If there is a $\hat{x}$ such that $\bar{F}(\hat{x}) < \bar{F}(\bar{x}) = \bar{\lambda}$, then $\tilde{\lambda} = \lambda_{min}(\tilde{C}) = \lambda_{min}(\sum \hat{x}_p S_p + Q) < \bar{\lambda}$.

Proof: $\hat{\lambda} = \sum \hat{x}_p \hat{y}' S_p \hat{y} + \hat{y}' Q \hat{y} \leqslant \sum \hat{x}_p \bar{y}' S_p \bar{y} + \bar{y}' Q \bar{y} < \sum \bar{x}_p \bar{y}' S_p \bar{y} + \bar{y}' Q \bar{y} = \bar{\lambda}$. Q.E.D.

In words: if the value of the completely linearized problem can be lowered, then λ_{min} can be lowered too. Theorem 3.11 gives us a set of necessary conditions for optimality, equivalent to the Fritz John conditions for this case. Unfortunately there are some quite bothersome complications. What happens, for example, if $\bar{x}$ is PSD. Then $\bar{\lambda} = 0$, and $\bar{y} = e$ is one of the minimising vectors. Now if we use e, then $e'S_p e = 0$ for all p, and we most certainly are in trouble. If $\bar{C} \neq nI - E$, then there is a $\bar{y}$ such that $\bar{C}\bar{y} = 0$ and $e'\bar{y} = 0$. We must always use that vector, never e. Another useful trick is switching from time to time to a linearization of the sum of the q smallest eigenvalues, with $1 < q < n-2$, and to use this linearization as $\bar{F}(x)$. This has the advantage that $\bar{F}(x)$, while remaining concave, assumes different values in the region where the matrices are PSD. The chances are that minimising this new $\bar{F}(x)$ brings us closer to the area where there really are negative eigenvalues.

Without these tricks the method runs something like this

i) Make the tableau.

ii) Find a basic feasible solution ($nI - E$ is always available, but it has the disadvantage that it lies deep in the PSD-region).

iii) Linearize the objective function and minimize it.

iv) If the solution we start iii) with is already the minimizing solution, then go to v), otherwise to iii).

v) If $\lambda_{min} < 0$ then make a new cut and go to i), otherwise stop.

Another refinement, which does not take too much time, is to investigate all neighboring solutions for every new basic feasible solution we find. There can not be more than $(k - 1)(l + 1)$ feasible neighbors. It may be profitable to insert such a step whenever it is called for. How easily the algorithm gets stuck, and how much time it takes on problems of the usual size ($n = 10$–20) is not known, and can hardly be guessed. The following 'branching' algorithm may also be useful. We start with $q = 1$.

i), ii) } As in previous method.

iii) Search for the neighboring basic feasible solution for which the sum of the q smallest eigenvalues is minimal.

iv) If the solution is the starting point itself and $q < n - 2$ then $q \rightarrow q + 1$, go to iii). If $q = n - 2$ then go to v). If the solution point and the starting point differ, then take the solution point as the new starting point, set $q = 1$, and go to iii).

v) As in previous method.

In this section we also mention a special technique which seems quite efficient for computing the minimum eigenvalue and the corresponding eigenvector.

Problem E: $F(x) = \frac{x'Ax}{x'x}$ min !

We use a variant of the cyclic-coordinate ascent method (Zangwill 1969 p 111-112 gives a convergence proof which also applies to this problem).

Subproblem E_i^k:

$$F_i^k(\theta) = \frac{(x + \theta e_i)'A(x + \theta e_i)}{(x + \theta e_i)'(x + \theta e_i)} = $$
$$= \frac{x'Ax + 2\theta x'a_{.i} + \theta^2 a_{ii}}{x'x + 2\theta x_i + \theta^2} \quad \text{min !}$$

Some algebra gives (supposing $x'x = 1$, and writing $x'Ax = \lambda$),

$$\frac{d\, F_i^k(\theta)}{d\,\theta} = 0 \quad \text{iff}$$

$$(x_i a_{ii} - x'a_{.i})\theta^2 + (a_{ii} - \lambda)\theta + (x'a_{.i} - \lambda x_i) = 0,$$

or

$$P\theta^2 + Q\theta + R = 0.$$

The minimizing θ is given by

$$\bar{\theta} = \frac{-Q + \sqrt{Q^2 - 4PR}}{2P}$$

if $P \neq 0$, otherwise $\bar{\theta} = 0$. The new x is defined as

$$x^+ = \bar{x} / (\bar{x}'\bar{x})^{\frac{1}{2}},$$

with

$$\bar{x} = x + \bar{\theta} e_i.$$

Then i becomes i + 1 until i = n (one cycle completed), and if the minimum of the absolute value of $\bar{\theta}$ in the cycle was larger then some small positive tolerance, then we start again with i = 1 (new cycle). The method was tested for the generalized eigen problem $\lambda = x'Ax / x'Bx$, and worked fast and reliable. Although we do not study the method in detail, the following theorem may be useful.

<u>Theorem 3.12</u>: If during a whole cycle $\bar{\theta} = 0$, then $Ax = \lambda x$. Moreover $Q^2 - 4PR \geq 0$.

Proof: Necessary for $\bar{\theta} = 0$ is that $R = 0$. If $R = 0$ for all i in a cycle then $Ax = \lambda x$. The second part is less obvious. We have the useful identity

$P + R = x_i Q$. Then $Q^2 - 4PR = Q^2 - 4P(x_i Q - P) = Q^2 - 4x_i PQ + 4P^2$. If $x_i PQ \leq 0$ the trivially $Q^2 - 4PR \geq 0$. If $x_i PQ > 0$, then $Q^2 - 4PR < 0$ implies $Q^2 + 4P^2 < 4x_i PQ = 2|x_i|\sqrt{4P^2Q^2}$. Consequently $Q^2 - 4PR < 0$ implies $(Q^2 + 4P^2)/2 < \sqrt{4P^2Q^2}$ which contradicts the arithmetic-geometric mean inequality. Q.E.D.

4 Selecting a particular solution

4.1 Optimal p-parsimoneous solutions

The first approach we shall discuss is looking for a parsimoneous solution. The alternative is looking for a representative solution. Both objectives sound desirable, but in most cases they yield far from identical results. This is intuitively clear: representativeness means that we want a vector somewhere deep in the interior of the cone, parsimony means that we want a vector on the edges. To be more exact: the optimal p-parsimoneous solution (in short: the O(p)P-solution) is defined as the C-matrix from the cone for which the sum of the p largest eigenvalues is a maximum (assuming that Tr(C) is some constant). The sum of the first p eigenvalues is a convex matrix function (cf Lemma 3.3), and the maximum of a convex function on a compact convex polyhedron occurs at one of the vertices. The solutions are thus given by those vertices for which $\sum_{i=1}^{p} \lambda_i$ is a maximum, it suffices to compute the spectra of all vertices. Combined with the 'complete' algorithm of section 3.4 this approach does not require separate special steps. Or: in the complete algorithm a relatively small amount of extra computation gives the O(p)P-solutions for $p=1,\dots,n-2$. Observe that the cutting off of $nI - E$ is especially important in this context. Adding $nI - E$ to a particular solution always makes is less parsimoneous for $p < n - 2$. Or, to put it differently, the O(p)P-solutions for $p \leqslant n-2$ are always in α-canonical form. It is obviously true that if there are (weak or semistrong) p-solutions then they are by definition also O(p)P-solutions.

The general problem can be written as

$$\sum_{i=1}^{p} \lambda_i(C) \quad \max !$$

$$\mathrm{Tr}(B_k C) \geqslant 0,$$

$$C \quad \text{SDC},$$

$$C \quad \text{PSD} \iff \lambda_n(C) \geqslant 0,$$

where the eigenvalues of C are of course ranked. These problems must essentially be solved separately for all different values of p. The O(p)P-solutions are not unique in general, which creates a complicated uniqueness problem for each p. The structure of the set S_p of all O(p)P-solutions is complicated. S_p is not convex: if $C_1, C_2 \in S_p$ then

$$\sum_{i=1}^{p} \lambda_i \left[\alpha C_1 + (1-\alpha) C_2 \right] \leq \alpha \sum_{i=1}^{p} \lambda_i(C_1) + (1-\alpha) \sum_{i=1}^{p} \lambda_i(C_2)$$

$$= \max_{C \in P} \sum_{i=1}^{p} \lambda_i(C).$$

Compare the discussion in section 2.1 and at the end of section 2.2.

There is an obvious motivation for maximizing the sum of the first p eigenvalues. We may be interested only in the the first p principal components, and we may be prepared to regard the remaining n - p eigen values as 'error'. In fact the procedure in this section is closely relatedto the familiar projection-procedures for NMS mentioned at the end of chapter I. If we translate their rationale into the space of C-matrices, then in this space we have a cone of matrices that satisfy the order relations in the data. Let V be this cone, and let W_p be the set of all C-matrices of rank $\leq$ p. In the projection-type algorithms we try to solve

$$\text{Problem PP:} \quad \min_{C_1 \in V} \; \min_{C_2 \in W_p} \| C_1 - C_2 \|^2,$$

where the norm is Euclidean. In a slightly different notation

$$\min_{C_1 \in V} D(C_1, W_p)^2,$$

D denoting the distance of point C_1 to set W_p, and, equivalently,

$$\min_{C_2 \in W_p} D(C_2, V)^2.$$

It is a well known result from least squares theory that

$$D(C_1, W_p)^2 = \sum_{i=p+1}^{n} \lambda_i^2(C_1),$$

and if we use the requirement that $Tr(C^2)$ is some constant, then problem PP is equivalent to

$$\max_{C_1 \in V} \sum_{i=1}^{p} \lambda_i^2(C_1),$$

which is very similar to the problem considered in this section. The main difference with the current projection-type algorithms is that they work in a different space (each distance defines an axis), and that they use norming of the objective function in order to circumvent degeneracy (in stead of restricting the attention to a particular polyhedron in the cone, as we do). A closer examination of the relation between the projection type algorithms and the methods described in this paper is given in chapter VI.

4.2 Minimum deviation solutions

Another interesting solution can be constructed if the δ_{ij} are numerical. Using

$$d_{ij} = \sqrt{c_{ii} + c_{jj} - 2c_{ij}},$$

we want

$$\| \delta_{ij} - d_{ij} \| \quad \text{min!}$$

$$\mathrm{Tr}(B_k C) \geq 0,$$

$$C \quad \text{SDC},$$

$$C \ \text{PSD}.$$

Unfortunately the function $\| \Delta - D \|$ is not convex. We may use the substitute

$$\| \delta_{ij}^2 - d_{ij}^2 \| = \| \delta_{ij}^2 - (c_{ii} + c_{jj} - 2c_{ij}) \| \quad \text{min!}$$

and this is convex. The problem can then be solved by any of the methods for convex programming (if we use the least squares norm, I would suggest as the most natural algorithm a combination of the cutting plane method and Beale's QP-method between the cuts). Now this least squares method may very well turn out to be a practical method for the NMS-problem. It is easily adapted to missing data and different vector norms. Although the method is related to the complete description method, it has a quite different rationale and implementation. The fact that we must square the δ_{ij} first may seem a disadvantage. If we do not square, however, we have a nonconvex problem and we are in for local minimum trouble.

A closely related problem was investigated by Hartley, Hocking, and Cooke (1967). Their objective was to solve the regression problem

$$Q(\beta) = \sum_{t=1}^{n} (y_t - E(y_t))^2 \qquad \text{min !}$$

where

$$E(y_t) = \beta_{00} + \beta' x_t + x_t' B x_t,$$

$$B = \{\beta_{ij}\} \qquad i,j = 1,\ldots,k .$$

$$\beta = \{\beta_{i0}\} \qquad i = 1,\ldots,k.$$

under the condition that B is PSD. As an algorithm they use the method of tangential approximation (Hartley and Hocking 1963). This method is quite similar to the cutting plane method, it is also based on linearization of both the objective function and the feasible region by, respectively, Wolfe's device and the generation of linear constraints when they are needed. It may be interesting to know that Hartley et al also give (both classical and Baysian) optimality properties of their estimates. They do this assuming normal theory, which is hardly tenable in our context.

We give an example that is somewhat different in nature from the other examples in this paper. Suppose that we have obtained the dissimilarities

$$\begin{array}{ccc} 0 & 6 & 4 \\ & 0 & 1 \\ & & 0 \end{array}$$

and we require $d_{12} > d_{13} > d_{23}$, but, moreover, we want the solution to be one-dimensional (by theorem 3.10 such a solution always exists, because $n = 3$). There are six possibilities for the order of x_1, x_2, and x_3 on the dimension. Each of these cases implies a certain partial order of the distances.

$$1:\ x_1 \geq x_2 \geq x_3 \Rightarrow d_{13} \geq d_{12} \wedge d_{13} \geq d_{23}.$$

$$2:\ x_1 \geq x_3 \geq x_2 \Rightarrow d_{12} \geq d_{13} \wedge d_{12} \geq d_{23}.$$

$$3:\ x_2 \geq x_1 \geq x_3 \Rightarrow d_{23} \geq d_{12} \wedge d_{23} \geq d_{13}.$$

$$4:\ x_2 \geq x_3 \geq x_1 \Rightarrow d_{12} \geq d_{23} \wedge d_{12} \geq d_{13}.$$

$$5:\ x_3 \geq x_2 \geq x_1 \Rightarrow d_{13} \geq d_{23} \wedge d_{13} \geq d_{12}.$$
$$6:\ x_3 \geq x_1 \geq x_2 \Rightarrow d_{23} \geq d_{13} \wedge d_{23} \geq d_{12}.$$

Only possibilities 2 and 4 are compatible with our inequality restrictions, they are 'mirror-images', so we may select either one of them. We take $x_2 \leq x_3 \leq x_1$, and thus

$$d_{12} = x_1 - x_2,$$
$$d_{13} = x_1 - x_3,$$
$$d_{23} = x_3 - x_2.$$

By this trick (which we shall call the Coombs-trick, because one-dimensional unfolding is based on it) we have reduced the system to one which only involves linear inequalities, and the complete description method becomes feasible without any iterations (and the least squares method without first squaring the dissimilarities). We are now working in the three-dimensional space R^3. The inequalities $d_{12} \geq d_{13} \geq d_{23}$; $d_{12}, d_{13}, d_{23} \geq 0$ define a (closed, convex, polyhedral) cone P (in fact they define two cones, if $x \in P$, then $-x$ also satisfies the inequalities). The solution set is $P \cup -P$.

D8 Compare this with our multidimensional approach: there we convexify by eliminating indeterminacy due to rotation, reflection, and translation. In the one-dimensional case there is no indeterminacy due to rotation, by not paying attention to $-P$ we eliminate the reflection-trouble, and the problem becomes convex. The fact that there is no rotational problem even makes the problem both linear and finite. The constraints are

$$\begin{array}{rcl}
d_{12} - d_{13} \geq 0 & \Leftrightarrow & -x_2 + x_3 \geq 0 \\
d_{13} - d_{23} \geq 0 & \Leftrightarrow & x_1 + x_2 - 2x_3 \geq 0 \\
d_{12} \geq 0 & \Leftrightarrow & x_1 - x_2 \geq 0 \\
d_{13} \geq 0 & \Leftrightarrow & x_1 - x_3 \geq 0 \\
d_{23} \geq 0 & \Leftrightarrow & -x_2 + x_3 \geq 0.
\end{array}$$

This system is redundant. If we eliminate the redundancies we obtain

$$\begin{array}{r}
-x_2 + x_3 \geq 0 \\
x_1 + x_2 - 2x_3 \geq 0.
\end{array}$$

The corresponding cone is not pointed (if $x_1 = x_2 = x_3$, then both inequalities are satisfied as equations). If we add the equation $x_1 + x_2 + x_3 = 0$ the cone becomes pointed (in a two-dimensional subspace), and at the same time we eliminate indeterminacy due to translation. There are two edges

$$(2 \ -1 \ -1) \qquad\qquad (1 \ -1 \ \ 0).$$

We shall refer to them as X_1 and X_2. Obviously $\rho^2_{min} = .75$, $|\rho|_{min} \simeq .866$. Observe that our multidimensional $r_{min} \leq .7071$, and our two one-dimensional edges are among the multidimensional ones (page 30, edge (2) and (3)). All solutions in P can be described by $\alpha\,(2 \ -1 \ -1) + \beta\,(1 \ -1 \ \ 0) =$ $(2\alpha + \beta \quad -\alpha - \beta \quad -\alpha)$, where α and β are two arbitrary nonnegative quantities. The distances are

$$d_{12} = 3\alpha + 2\beta\ ,$$
$$d_{13} = 3\alpha + \beta\ ,$$
$$d_{23} = \beta\ ,$$

and we must minimize

$$\begin{aligned} F &= (3\alpha + 2\beta - 6)^2 + (3\alpha + \beta - 4)^2 + (\beta - 1)^2 = \\ &= 9\alpha^2 + 4\beta^2 + 12\alpha\beta - 36\alpha - 24\beta + 36 + \\ &\quad\ 9\alpha^2 + \beta^2 + 6\alpha\beta - 24\alpha - 8\beta + 16 + \\ &\quad\ \beta^2 - 2\beta + 1 = \\ &= 18\alpha^2 + 6\beta^2 + 18\alpha\beta - 60\alpha - 34\beta + 53. \end{aligned}$$

$$\frac{\partial F}{\partial \alpha} = 36\alpha + 18\beta - 60.$$

$$\frac{\partial F}{\partial \beta} = 18\alpha + 12\beta - 34.$$

We must solve the linear system

$$\begin{bmatrix} 36 & 18 \\ 18 & 12 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 60 \\ 34 \end{bmatrix}.$$

The solution is $\bar{\alpha} = 1$, $\bar{\beta} = \frac{4}{3}$. The least squares solution is thus $\frac{1}{3}\,(10 \ -7 \ -3)$, with distances $d_{12} = 5.66..$; $d_{13} = 4.33..$; $d_{23} = 1.33..$ The question how representative this solution is can now be formulated as follows: find the solution y with minimum squared correlation coefficient with the least squares solution x, i.e. $\rho^2(x,y) = \min_{z \in P} \rho^2(x,z)$.

It can be proved that the solution vector is an edge of P (cf section 4.4). The squared correlations of X_3 with the edges are 150/158 and 289/316, or $|\rho|_{min} = .956$. The least squares solution X_3 is thus considerably more representative then either X_1 or X_2.

If we apply the 'metric' method outlined by Torgerson, and construct the matrix of 'pseudo-scalar-products' (cf Appendix A) from the squared dissimilarities we obtain a matrix that is not PSD (this follows, of course, from the fact that our dissimilarities violate the triangular inequality). The ratio $-\lambda_1/\lambda_2$ is, however, about 15, so the reproduction will still be quite perfect. The first dimension of the metric solution correlates .999 with the least squares solution, and its $|\rho|_{min}$ is .955.

4.3 Degrees of faithfulness

A different way to find more unique solutions is to construct C-matrices that are 'faithful to degree n'. Fundamentally this concept is due to Tversky (1964).

Definition 4.1: A solution is said to be (weakly, semi-strongly, strongly) faithfull of degree n iff the solution preserves (weakly, semi-strongly, strongly) the order-relations between all n^{th}-order differences of squared dissimilarities.

Thus an ordinary solution (as we have considered them) is faithful of degree zero. In the strong case

$$\text{Degree 0:}\quad \delta_{ij}^2 \gtrless \delta_{kl}^2 \iff d_{ij}^2 \gtrless d_{kl}^2,$$
$$\text{Degree 1:}\quad \delta_{ij}^2 - \delta_{kl}^2 \gtrless \delta_{i'j'}^2 - \delta_{k'l'}^2 \iff$$
$$d_{ij}^2 - d_{kl}^2 \gtrless d_{i'j'}^2 - {}_{k'l'}^2.$$

Observe that the first is a special case of the second (set k = k', l = l'), and so on. Again: the complete description method cannot be used unless we square the dissimilarities first, because it is not true that

$$d_{ij} - d_{kl} \gtrless d_{i'j'} - d_{k'l'} \iff$$
$$d_{ij}^2 - d_{kl}^2 \gtrless d_{i'j'}^2 - d_{k'l'}^2.$$

This obviously means that we must have numerical dissimilarities, although the definition of a faithful solution could easily be made to depend on order relations only. In the one-dimensional case there is no such disadvantage.

If P_k denotes the convex cone of solutions that are faithful to degree k, then

$$P_0 \supseteq P_1 \supseteq P_2 \supseteq \;\ldots\ldots$$

If k is large, and the system has a faithful solution of degree k, then we expect this solution to be almost unique. We may be interested (as Tversky is in his discussions of faithful solutions for additive conjoint measurement) in the maximal k for which the system is still solvable (a solution of this representation theorem and some information about the uniqueness problem is given in appendix B).

We continue our one-dimensional example, in which we have the first order information $\delta_{12} - \delta_{13} = 2$; $\delta_{13} - \delta_{23} = 3$. Thus

$$\delta_{13} - \delta_{23} > \delta_{12} - \delta_{13},$$

and we require in addition that

$$(x_1 - x_3) - (x_3 - x_2) - (x_1 - x_2) + (x_1 - x_3) =$$
$$x_1 + 2x_2 - 3x_3 \geqslant 0.$$

Let $b' = (1 \;\; 2 \; -3)$. Our edges were $X_1' = (2 \; -1 \; -1)$, and $X_2' = (1 \; -1 \;\; 0)$. The first one satisfies the new restraint ($b'X_1 = 3$), the second one does not ($b'X_2 = -1$). The new edges are

$$X_1' = (2 \; -1 \; -1),$$
$$X_2' = (5 \; -4 \; -1).$$

Now $\rho^2(X_1,X_4) = 25/28$, $|\rho|_{min} = .945$. For the least squares solution X_3 we obtain $b'X_3 = 5/3$, which means that X_3 is in the cone P_1 and that it is still the least squares solution.

<u>Definition 4.2</u>: A linear functional of the δ_{ij}^2 will be called an <u>integer-contrast of order k</u> iff the coefficients a_{ij} in $\sum_i \sum_j a_{ij} \delta_{ij}^2$ are such that for all i,j

i) a_{ij} is a (positive, negative, or zero) integer.

ii) $-k \leqslant a_{ij} \leqslant +k$.

iii) $\sum_i \sum_j a_{ij} = 0$.

In order to obtain more and more unique solutions we may use constraints of the type

$$\sum\sum a_{ij}\delta^2_{ij} \geqslant 0 \Leftrightarrow \sum\sum a_{ij}d^2_{ij} \geqslant 0$$

(linear-k constraints). The constraints used to define P_0 are linear-1 constraints, the added constraint in our last example is a linear-2 constraint (in the d_{ij}, it turns out to be a linear-3 constraint in the x_i). If we choose $a_{12} = 2$, $a_{13} = -3$, and $a_{23} = 1$ (a linear-3 constraint), then $\sum\sum a_{ij}\delta_{ij} = 12 - 12 + 1 = 1 > 0$, so

$$2(x_1 - x_2) - 3(x_1 - x_3) + (x_3 - x_2) =$$
$$-x_1 - 3x_2 + 4x_3 \geqslant 0$$

is the corresponding requirement. Let $c' = (-1\ -3\ \ 4)$. Applying this to the edges X_1 and X_4 we obtain $c'X_1 = -3$ and $c'X_4 = 3$. The new edges are

$$X'_4 = (\ 5\ -4\ -1),$$
$$X'_5 = (\ 7\ -5\ -2),$$

and ρ^2_{min} is now up to 3249/3276. For all practical purposes we are working on an interval scale (a ratio scale of distances).

If we work with numerical δ_{ij} we can sketch the real function F that relates δ_{ij} to d_{ij}. Requiring strong faithfulness of degree zero means requiring F to be strictly monotonic increasing (some people would prefer to talk of a scattergram in stead of a function, Guttman has suggested the name Shepard-diagram). Suppose that we require in addition strong monotonicity of the positive first-order differences . In words this means that we want small changes in the argument of F to correspond with small changes in the value of F, i.e. we want F to be smooth. There is an obvious connection with the work of Shepard and Carroll (1966) on maximizing continuity (but I do not want to use that word in this context, cf also Kruskal and

Carroll (1969)). It is not too difficult to require both monotonicity and smoothness, it is more complicated to require smoothness without monotonicity. One approach is the one outlined by Shepard, Carroll, and Kruskal. An D9. essentially nonmetric alternative is outlined in appendix E.

The effect of adding constraints that make F more smooth is illustrated in figure 5 a-c. The distances using edges X_1 and X_2 are

$$d_{12} = 3\alpha + 2\beta,$$
$$d_{13} = 3\alpha + \beta,$$
$$d_{23} = \beta.$$

We require $d_{12} + d_{13} + d_{23} = \delta_{12} + \delta_{13} + \delta_{23} = 11$, and eliminate β. This gives

$$d_{12} = \frac{11}{2},$$
$$d_{13} = \frac{6\alpha + 11}{4},$$
$$d_{23} = \frac{11 - 6\alpha}{4},$$

with $0 \leq \alpha \leq \frac{11}{6}$. For the edges X_1 and X_4 we obtain in a similar way

$$d_{12} = \frac{11}{2},$$
$$d_{13} = \frac{3\alpha + 11}{4},$$
$$d_{23} = \frac{11 - 6\alpha}{4},$$

with $0 \leq \alpha \leq \frac{11}{6}$, and for X_4 and X_5

$$d_{12} = \frac{11}{2},$$
$$d_{13} = \frac{33 - 6\alpha}{8},$$
$$d_{23} = \frac{6\alpha + 11}{8},$$

with $0 \leq \alpha \leq \frac{11}{18}$. The small circles in figure 5 a-c give the value of d for $\alpha = 0$, the crosses for $\alpha = \frac{11}{6}, \frac{11}{6}, \frac{11}{18}$ respectively.

These plots show that F becomes not only less variable and more smooth, in fact F becomes linear (cf Appendix B for an explanation). This result may be useful too in projection-type programs, where it is hardly any extra trouble to require monotonicity of positive first order differen-

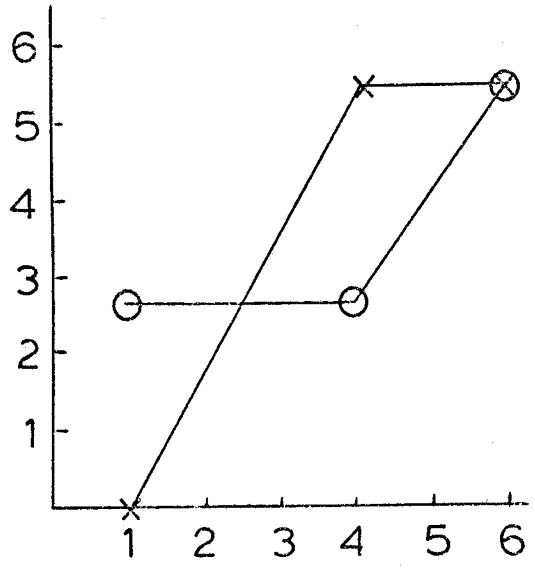


fig 5a
crosses $\alpha = 11/6$
circles $\alpha = 0$

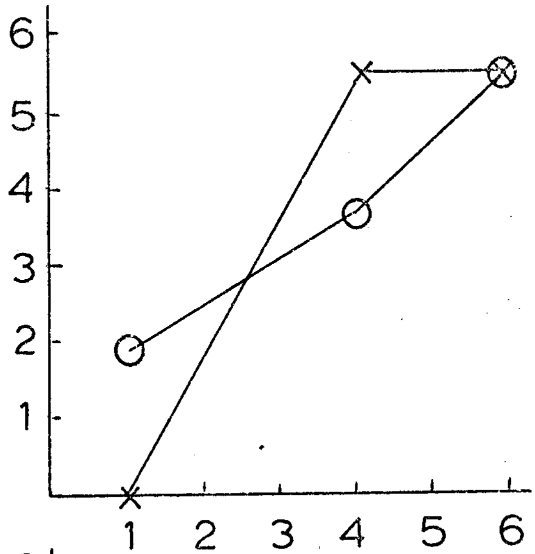


fig 5b
crosses $\alpha = 11/6$
circles $\alpha = 0$

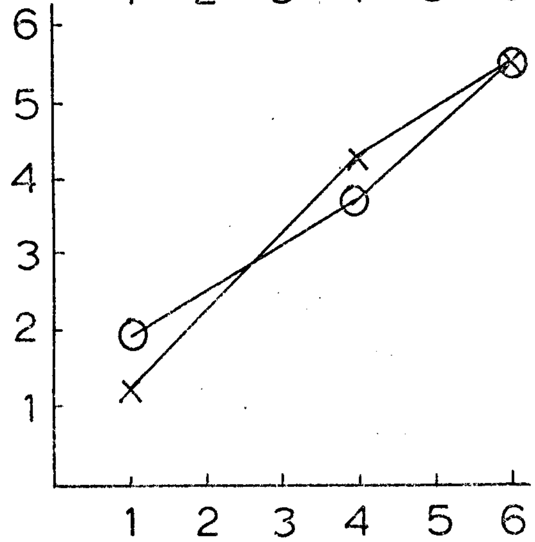


fig 5c
crosses $\alpha = 11/18$
circles $\alpha = 0$

ces. We may even include inequalities of the type $d_{ij} \geq d_{kl} - d_{pq}$ (≥ 0) and obtain a joined weak order of dissimilarities and differences. This will tend to produce less jagged Shepard-diagrams, and possibly also fewer degenerate solutions.

4.4 Representative solutions

There are two general methods available for finding a representative solution in a polyhedral convex cone. The first one is due to Frank Goode of the University of Michigan. His procedures are unpublished but Coombs describes them in 1964 p 96-102 for the one-dimensional unfolding problem and in 1966 p 102-107 for additive conjoint measurement. Goode's method is essentially a convenient 'work-sheet'-type method to compute the edges of the cone, and his 'equal-Δ'-solution is simply the average of these edges. There are some arbitrary elements in this method, due to the fact that edges are determined only up to multiplication with a positive constant. Goode chooses them in such a way that they consist entirely of integers, and that there is no integer edges with smaller length. The method mentioned just before theorem 3.2 is a constructive proof that integer edges always exist, by requiring that the coordinate values have no common divisor other than one we obviously make them unique. We shall call these edges in Goode-shape, and make some additional comments on Goode's equal-Δ method and its applicability after a consideration of the alternative method due to Abelson and Tukey (A & T for short).

The basic theory of this method is developed in the paper A & T (1963), a number of computational examples is collected in A & T (1959). The idea is simple. Given any solution C_1 in the pointed polyhedral convex cone P, we can compute the solution C_2 such that

$$r(C_1, C_2) = \min_{C \in P} r(C_1, C).$$

This is a measure of the representativeness of C_1, and we have already used it thus in our LS-example on page 60-61. The proposal of A & T is to select C_1 in such a way that this minimum value of r is maximized. This we shall problem AT:

$$\max_{C_1 \in P} \ \min_{C_2 \in P} \ r(C_1, C_2).$$

The maximin solution of r is written as r_{AT}; if we are talking loosely about the AT-solution we do not mean the pair (C_1, C_2) which solves AT, but we mean C_1 from this pair. In the theoretical development of the AT-procedure , we shall use different techniques as A & T do. This is mainly because our method shows the algorithmic implications more clearly, but also to preserve continuity with the rest of the paper (we also use somewhat different terminology and a different notation). We let $S = \{S_1, \ldots, S_p\}$ denote the set of edges of P, scaled in such a way that $Tr(S_k^2) = 1$.

Theorem 4.1: If $(\bar{C}_1, \bar{C}_2)$ solves problem AT, then $\bar{C}_2 \in S$.

Proof: This is half of theorem 2.7, and the proof is almost identical. $\bar{C}_2$ can be written as $\sum \alpha_k S_k$, with $\alpha_k \geqslant 0$. Consequently for fixed C_1 (of unit lenght) the problem

$$\min_{C_2 \in P} \ r(C_1, C_2)$$

is equivalent to

$$\min_{\alpha_k \geqslant 0} \frac{\sum \alpha_k Tr(C_1 S_k)}{\sqrt{\sum \alpha_k \alpha_l Tr(S_k S_l)}}.$$

As in the proof of theorem 2.7 we have

$$\sum \alpha_k Tr(C_1 S_k) \geqslant \sum \alpha_k \min_{k=1}^{p} Tr(C_1 S_k),$$

$$\sum \alpha_k \alpha_l Tr(S_k S_l) \leqslant (\sum \alpha_k)^2 \max_{k,l} Tr(S_k S_l) = (\sum \alpha_k)^2,$$

and

$$r(C_1, C_2) \geqslant \min_{k=1}^{p} r(C_1, S_k)$$

for all $C_2 \in P$, with equality iff C_1 is the edge which actually minimizes $r(C_1, S_k)$ over S. Q.E.D.

This theorem is essentially a formalization of A & T 1963 p 1352-1353. They consider the slightly more general case in which there may be vectors in the cone with $r = -1$ (which implies that the cone is not pointed). Even in this more general case (any nondegenerate cone) the worst possible outcome is $r_{AT} = 0$. If P is the half-space $\{C \mid Tr(BC) \not\leq 0\}$, then C_1 may be taken as $-B$ and C_2 as any vector in the hyperplane $Tr(BC) = 0$. If P is the whole space (a degenerate cone) then $r_{AT} = -1$: for each C_1 we have $\min_{C_2 \in P} r(C_1, C_2) = -1$. At the other extreme: if P is a single ray then $r_{AT} = 1$.

Consider the following p problems Q_k.

$$\max_{C \in P} \ Tr(CS_k).$$

$$Tr(CS_k) \leq Tr(CS_l), \qquad (l=1,\ldots,p)$$

$$Tr(C^2) \leq 1.$$

Let the solutions be $\bar{C}^k$ with value $\bar{r}^k$ (obviously $\bar{r}^k \geq 0$ for all k). It is easy to see that $\max_{k=1}^{p} \bar{r}^k = r_{AT}$, or: the AT-solution is the solution $\bar{C}^k$ with the highest value of $\bar{r}^k$. The algorithm (solving all problems Q_k) suggested by this result is quite inefficient (as we shall see).

Consider now the problem Q'_k which is identical to Q_k, but without the constraints $C \in P$. The solutions of Q'_k are denoted by $\bar{C}'_k$ with value $\bar{r}'_k$. Let $\bar{C}'_q$ be the solution with the highest value of $\bar{r}'_k$.

<u>Lemma 4.1</u>: A feasible point $\bar{C}'_k$ solves Q'_k iff there are nonnegative multipliers $\beta^k, \mu^k_1, \ldots, \mu^k_p$ such that

$$KT^k_1: \quad \beta^k \bar{C}'_k = (1 - \sum \mu^k_l) S_k + \sum \mu^k_l S_l.$$

$$KT^k_2: \quad (\forall\, l) : \mu^k_l (Tr(\bar{C}'_k S_k) - Tr(\bar{C}'_k S_l)) = 0.$$

$$KT^k_3: \quad \beta^k (Tr(\bar{C}'^{\,2}_k) - 1) = 0.$$

Proof: These are simply the Kuhn-Tucker conditions for the convex programming problem Q'_k. People who are not familiar with these conditions can look them up in Kuhn and Tucker (1951) and, for that matter, in any textbook on NLP. An especially thorough treatment can be found in the recent book of Mangasarian (1969), and a very general one in the contribution of Hurwitz to the book Arrow et al (1958). A short, systematic, precise and very readible discussion is included in the second edition of Berge's excellent book (Berge, 1966). Q.E.D.

Corollary 4.1: $\bar{C}'_q$ solves Q'_q iff KT^q_1, KT^q_2, KT^q_3, and

$$(\forall l): Tr(\bar{C}'_q S_q) \geqslant Tr(\bar{C}'_l S_l).$$

Lemma 4.2: If $\bar{C}'_q$ solves Q'_q then there is a subset $K \subset I_p = \{1,\ldots,p\}$ such that

$$KQ1: \bar{C}'_q = \sum_{l \in K} \bar{\alpha}_l S_l, \text{ with } (\forall_K l): \bar{\alpha}_l > 0.$$

$$KQ2: (\forall_K l): Tr(\bar{C}'_q S_l) = \bar{r}'_q.$$

$$KQ3: (\forall_{I_p - K} l): Tr(\bar{C}'_q S_l) \geqslant \bar{r}'_q.$$

$$KQ4: Tr(\bar{C}'^{\,2}_q) = 1.$$

Proof: $Tr(\bar{C}'_q S_q) = \bar{r}'_q = \max_{k \in I_p} \bar{r}'_k \geqslant \max_{k \in I_p} \bar{r}^k = r_{AT} > 0$. If $Tr(\bar{C}'^{\,2}_q) = \delta < 1$, then $\delta^{-\frac{1}{2}}\bar{C}'_q$ would be feasible for Q'_q and give a value for the problem of $\delta^{-\frac{1}{2}}\bar{r}'_q > \bar{r}'_q$, which contradicts the fact that $\bar{C}'_q$ solves Q'_q. This proves KQ4. KT^q_2 can be written as

$$Tr(\bar{C}'_q S_q) < Tr(\bar{C}'_q S_l) \Rightarrow \mu^q_l = 0,$$

$$\mu^q_l > 0 \Rightarrow Tr(\bar{C}'_q S_q) = Tr(\bar{C}'_q S_l).$$

Define $K = \{l \mid \mu^q_l > 0\}$. Then KQ2 and KQ3 are automatically true. It follows from KT1-KT3 that we may always assume $q \in K$. We may write KT^q_1 as

$$\beta^q \bar{C}'_q = S_q - \sum \mu^q_l (S_q - S_l).$$

Multiplying both sides with $\bar{C}'_q$, taking traces, and using KQ4 and KT^q_2 proves that $\beta^q = \bar{r}'_q > 0$. Suppose $K = \{q\}$. Then KT^q_1 gives $\beta^q \bar{C}'_q = S_q$, and $(\beta^q)^2 = (\bar{r}'_q)^2 = 1$, $\bar{C}'_q = S_q$, and P is the ray S_q. In this trivial case $\bar{C}'_q$ is, of course, the AT-solution. We exclude this case by assuming $\{q\} \subset\subset K$. If $p, q \in K$, then $\bar{C}'_q$ is also feasible for Q'_p, and $Tr(\bar{C}'_q S_q) = Tr(\bar{C}'_q S_p) \leqslant Tr(\bar{C}'_p S_p)$. Using corollary 4.1 this implies $Tr(\bar{C}'_q S_q) = Tr(\bar{C}'_p S_p) = Tr(\bar{C}'_q S_p)$, and $\bar{C}'_q$ solves both Q'_p and Q'_q. In fact it solves Q'_l for all $l \in K$.

We have

$$(\forall_K l): \bar{r}'_q \bar{C}'_q = \sum_{k \in K} \bar{\mu}^l_k S_k,$$

with

$$\bar{\mu}^l_k = \begin{cases} \mu^l_k & \text{if } k \neq l, \\ 1 - \sum_{k \in K - \{l\}} \mu^l_k & \text{if } k = l. \end{cases}$$

Moreover,

$$(\forall_K l): \operatorname{Tr}(\overline{C}'_q S_l) = \overline{r}'_q.$$

Consider an arbitrary vector $\overline{C}$ in the subspace spanned by the S_l ($l \in K$) such that $\operatorname{Tr}(\overline{C}^2) = 1$, and $(\forall_K l)$: $\operatorname{Tr}(\overline{C}S_l) = \delta$. Let $\overline{C} = \sum x_l S_l$, and suppose $\delta = 0$. Then $1 = \operatorname{Tr}(\overline{C}^2) = \sum x_l \operatorname{Tr}(\overline{C}S_l) = \delta \sum x_l = 0$, which is impossible. Thus $\delta \neq 0$. Let $C^+ = \delta^{-1}\overline{C}$, and $C^+_q = (\overline{r}'_q)^{-1}\overline{C}'_q$. Then $(\forall_K l)$: $\operatorname{Tr}(C^+S_l) = \operatorname{Tr}(C^+_q S_l) = 1$, or $\operatorname{Tr}(S_l(C^+ - C^+_q)) = 0$. The vector $C^+ - C^+_q$ is also a linear combination of the S_l, say with coefficients γ_l. Thus $\operatorname{Tr}(S_l \sum \gamma_k S_k) = \sum \gamma_k \operatorname{Tr}(S_k S_l) = 0$ for all l, which implies that $\sum \gamma_k S_k = 0$, or $C^+ = C^+_q$, or $\delta = \pm \overline{r}'_q$. The only possible $\overline{C}$ different from $\overline{C}'_q$ is thus $-\overline{C}'_q$ (this is lemma C of A & T, p 1367). It follows that $(\forall_K l)(\forall_K k)$: $\overline{\mu}^l_k = \overline{\mu}^q_k$, and in particular it follows that $\overline{\mu}^k_k = \overline{\mu}^q_k = \mu^q_k > 0$ and $\overline{\mu}^q_q = \overline{\mu}^k_q = \mu^k_q > 0$ for $k \neq q \in K$. Let $\overline{\alpha}_l = (\overline{r}'_q)^{-1}\overline{\mu}^q_l$, then KQ1 is true. Q.E.D.

<u>Lemma 4.3</u>: If there is a feasible point C'_q and a subset $L \subset I_p$ such that KQ1-KQ4 are true for L, C'_q, r'_q, and α_l, then C'_q solves Q'_q.

Proof: Suppose there is a $C^+_q \neq C'_q$ that solves Q'_q. Then $(\forall_L l)$: $\operatorname{Tr}(C^+_q S_l) > \operatorname{Tr}(C'_q S_l) = r'_q$. This implies $\sum \alpha_l \operatorname{Tr}(C^+_q S_l) = \operatorname{Tr}(C'_q C^+_q) > \sum \alpha_l \operatorname{Tr}(C'_q S_l) = \operatorname{Tr}(C'^{\,2}_q) = 1$, which is impossible. (This is A & T's sufficiency argument on page 1368). Q.E.D.

<u>Corollary 4.2</u>: $\overline{C}'_k$ solves Q'_k but not Q_k iff $\overline{r}'_k > \overline{r}^k$ iff $\overline{\mu}^k_k < 0$ iff $\sum_{l \in K} \overline{\mu}^k_l > 1$ iff $k \notin K$.

<u>Lemma 4.4</u>: $\overline{C}'_q = C_{AT}$.

Proof: If $r_{AT} \geq 0$, then $\operatorname{Tr}(C^2_{AT}) = 1$. If $C_{AT} \neq \overline{C}'_q$ then we get a contradiction as in lemma 4.3. Q.E.D.

<u>Lemma 4.5</u>: If $\overline{r}'_q > 0$, then the solution of Q'_q is unique.

Proof: Suppose ${}_1\overline{C}'_q$ and ${}_2\overline{C}'_q$ are two different optimal solutions of Q'_q, and let ${}_3\overline{C}'_q = \delta\,{}_1\overline{C}'_q + (1-\delta)\,{}_2\overline{C}'_q$, with $0 < \delta < 1$. Then $\operatorname{Tr}({}_3\overline{C}'_q S_l) = \overline{r}'_q$ for $l \in K$, and $\operatorname{Tr}\,({}_3\overline{C}'_q)^2 = \delta^2 + (1-\delta)^2 + 2\delta(1-\delta)\operatorname{Tr}({}_1\overline{C}'_q\,{}_2\overline{C}'_q) < 1$. Consequently there is a $\gamma > 1$ such that $\gamma\,{}_3\overline{C}'_q$ is still feasible, but $\operatorname{Tr}(S_l(\gamma\,{}_3\overline{C}'_q)) =$

$= \int \bar{r}'_q > \bar{r}'_q$ which contradicts the fact that ${}_1\bar{C}'_q$ and ${}_2\bar{C}'_q$ are optimal. Q.E.D.

We are now ready to state the key-theorem of Abelson and Tukey. In our proofs we did in fact never use the fact that P is polyhedral, it may as well have an infinite number of edges. We did use the fact that every element of P can be written as a positive linear combination of a finite number of edges, and this is true in any finite-dimensional space.

Theorem 4.2: If $\mathcal{S}$ is the set of edges of a pointed cone P in a real finite-dimensional linear space, then the triple $(\bar{r}, C_1, C_2)$ solves the problem

$$\bar{r} = \max_{C_1 \in P} \min_{C_2 \in P} r(C_1, C_2)$$

iff there is a finite set $\mathcal{S}' \subset \mathcal{S}$ (indexed by K) such that

$$AT1: \; C_1 = \sum_{l \in K} \alpha_l S_l \text{ with } (\forall_K l): \; \alpha_l > 0,$$

$$AT2: \; (\forall_K l): \; Tr(C_1 S_l) = \bar{r} > 0,$$

$$AT3: \; (\forall_{\mathcal{S}-\mathcal{S}'} S): \; Tr(C_1 S) \geqslant \bar{r}.$$

C_1 is unique, C_2 is any one of the S_l, $l \in K$.

This theorem (or rather a slightly less general version of it) is proved by A & T on p 1366-1368, but their proof differs considerably from ours. They use more elementary and geometrical reasoning and not the problems Q_k and Q'_k with their KT-conditions. As a matter of fact they derive the KT-conditions for AT all over again, in stead of applying them directly. We have chosen this route, because it has important algorithmic implications, and because we assume that our readers are familiar with the basic facts from NLP-theory. The most difficult part of the proof is the necessity of KQ1. In their version of the proof A & T use the separation theorem for convex polyhedral sets and cones (the most useful version is Goldman (1956) p 50). This is, of course, a more fundamental result than the KT-theorem, although it will be familiar to the same circle of readers. I think it is fair to say that the method of proof is closely related to the algorithmic methods I intend to use. The same thing is true for A & T.

A & T sketch some heuristic iterative approaches based on judicious trial-and-error and on making all possible use of the special structure of the particular problem. They also mention the possibility of an exhaustive approach (investigate all subsets of edges) and conclude, correctly, that this is most certainly inpractical. And they mention the possibility of a QP-approac but do not work it out because their examples 'yield with satisfactory ease to much less powerful tools' (1963, p 1356). In any case they assume that the edges of P are known beforehand (and, a fortiori, that there is only a finite number of them). This assumption can be relaxed to the weaker assumption that a finite spanning set of P is known, which may be important.

In the EMMS-application there are three main problems. In the first place the number of edges of the cone P may be infinite. This is not too important, however, given the complete method. We have a number of polyhedral cones $Q_0 \supset Q_1 \supset \dots$ which all contain P, and the AT-solutions for $Q_0, Q_1, \dots$ converge to the AT-solution for P. In the second place the number of elements in the spanning sets for Q_s is usually very large. This makes the heuristic methods of A & T quite impractical. We must use QP, the more powerful tool. A detailed analysis of the problem shows, I think, that the QP-approach is computationally very simple.

Again we consider the case in which we do know all edges. Problem Q_k is equivalent to

$$\sum_{l=1}^{p} \lambda_l r_{kl} \quad \text{max !}$$

$$\sum \lambda_l (r_{kl} - r_{tl}) \qquad 0, \qquad\qquad (t=1,\dots,p)$$

$$\sum\sum \lambda_l \lambda_q r_{lq} \leq 1,$$

$$\lambda_l \geq 0, \qquad\qquad (l=1,\dots,p)$$

with $r_{lq} = Tr(S_l S_q)$. Define $\bar{R}^k$ by $\bar{r}^k_{lq} = r_{kl} - r_{ql}$. Then Q_k can be written as

$$\lambda' r_{k.} \quad \max !$$
$$\bar{R}^k \lambda \leqslant 0,$$
$$\lambda' R \lambda \leqslant 1,$$
$$\lambda \geqslant 0.$$

In Q_k' we drop the requirement $\lambda \geqslant 0$ (the solution is not required to lie in the cone, only in the subspace spanned by the S_1). The results of theorem 4.2 can now be formulated as: if $\bar{\lambda}^k$ solves the modified form of Q_k' with corresponding maximum value $\bar{r}_k'$ and subset $S^k = \left\{ S_1 \middle| \bar{\lambda}^{k\prime} r_{1.} = \bar{r}_k' \right\}$, then $C_1 = \sum \bar{\lambda}_1^k S_1$ and $C_2 \in S^k$ arbitrary solve AT iff $\bar{\lambda}_k^k > 0$. This is, of course, a very convenient criterion to decide whether a particular solution is optimal or not.

The problems Q_k' are not yet as simple as they could be. The following theorem helps to simplify them further. A slightly different version was proved by Zoutendijk (1960 p 81) in the quite different context of finding normalized 'feasible directions' in the steps of iterative NLP-procedures. Consider problem Q_k^2

$$\lambda' R \lambda \quad \min !$$
$$\bar{R}_k \lambda \leqslant 0,$$
$$\lambda' r_{k.} = 1.$$

<u>Theorem 4.3</u>: If $\bar{\lambda}^k$ solves Q_k' and $\bar{r}_k' > 0$, then for some $\gamma > 0$ $\bar{\lambda}_+^k = \gamma \bar{\lambda}^k$ will solve Q_k^2.

Proof: $\bar{r}_k' = r_{k.}' \bar{\lambda}^k > 0$. Set $\bar{\gamma} = (\bar{r}_k')^{-1}$, and suppose $\bar{\lambda}_+^k$ solves Q_k^2. Clearly $\bar{\gamma} \bar{\lambda}^k$ is also feasible for Q_k^2, and thus $\bar{\lambda}_+^{k\prime} R \bar{\lambda}_+^k \leqslant \bar{\gamma}^2 \bar{\lambda}^{k\prime} R \bar{\lambda}^k = \bar{\gamma}^2$. Suppose $\bar{\lambda}_+^{k\prime} R \bar{\lambda}_+^k = \delta < \bar{\gamma}^2$. If $\delta = 0$ then $R \bar{\lambda}_+^k = 0$, and $r_{k.}' \bar{\lambda}_+^k = 0$, which means that $\bar{\lambda}_+^k$ would not be feasible for Q_k^2. Con-se-quently $\delta > 0$ and $\delta^{-\frac{1}{2}} \bar{\lambda}_+^k$ is feasible for Q_k', and $\delta^{-\frac{1}{2}} r_{k.}' \bar{\lambda}_+^k > \bar{\gamma}^{-1} r_{k.}' \bar{\lambda}_+^k = r_{k.}' \bar{\lambda}^k$, which contradicts the fact that $\bar{\lambda}^k$ is optimal for Q_k'. Consequently $\delta = \bar{\gamma}^2$, and $\bar{\gamma} \bar{\lambda}^k$ solves Q_k^2. Q.E.D.

We shall now eliminate λ_k from the problem Q_k^2 by using $\lambda' r_{k.} = \lambda_k + \sum_{1 \neq k} r_{kl} \lambda_1 = 1$. Then

$$\lambda' R \lambda = 1 + \sum_{s,t \neq k} \sum (r_{st} - r_{ks} r_{kt}) \lambda_s \lambda_t,$$

and

$$\sum \lambda_l (r_{kl} - r_{tl}) = (1 - r_{tk}) - \sum_{l \neq k} (r_{tl} - r_{tk} r_{kl}) \lambda_l.$$

Define $\bar{R}^k_+$ with $\bar{r}^k_{+ij} = r_{ij} - r_{ik} r_{jk}$ and $\bar{r}^k_+$ with $\bar{r}^k_{+i} = 1 - r_{ik}$. Evidently $\bar{r}^k_+(i,k) = \bar{r}^k_+(k,j) = \bar{r}^k_+(k) = 0$ for all i,j. Construct S^k of order $p-1$ by leaving out row and column k from $\bar{R}^k_+$ and s^k by leaving out element k from $\bar{r}^k_+$. Then Q^2_k can be reformulated as the equivalent problem Q^3_k:

$$1 + \lambda' S^k \lambda \quad \text{min!}$$

$$S^k \lambda \geq s^k.$$

This is a very simple QP-problem but still not simple enough for our taste. It could be simplified further by using a transformation that diagonalizes S^k, but we shall not do this. Other simplifications are available, which seem more effective.

<u>Theorem 4.4</u>: S^k is PSD for all k, if R is PD then S^k is PD for all k, $s^k \geq 0$ for all k.

Proof: It clearly suffices to prove this for $k = 1$. R can be partitioned as

$$R = \begin{bmatrix} 1 & r' \\ r & R \end{bmatrix}$$

and $S^1 = R - rr'$. Because R is PSD we have for all real α and all $p-1$ element vectors $z \neq 0$ that $\alpha^2 + 2\alpha r'z + z'Rz \geq 0$, which implies that the discriminant $4(r'z)^2 - 4z'Rz \leq 0$ for all $z \neq 0$. But $z'S^1z = z'Rz - (z'r)^2$, and S^1 is PSD. If R is PD, then $z'Rz - (z'r)^2 = z'S^1z > 0$ for all $z \neq 0$, and S is PD. By definition $s^1_i = 1 - r_{1i} \geq 0$ for all i. Q.E.D.

Consider the problem Q^4_k: find vectors x and y such that

$$S^k x - y = s^k,$$

$$x \geq 0 \ , \ y \geq 0 \ , \ x'y = 0.$$

<u>Theorem 4.5</u>: If $(\bar{x},\bar{y})$ solves Q^4_k, then $\bar{x}$ solves Q^3_k. If $\bar{x}$ solves Q^3_k and $\bar{r}'_k > 0$, then $(\bar{x}, S^k\bar{x} - s^k)$ solves Q^4_k.

Proof: Q_k^3 is a convex programming problem. The KT-conditions (necessary and sufficient) for a feasible $\bar{x}$ to be a solution is that there are multipliers $\bar{z}_1, \ldots, \bar{z}_{p-1} \geqslant 0$ such that

$$S^k\bar{x} = S^k\bar{z} ,$$

$$\bar{z}'(S^k\bar{x} - s^k) = 0.$$

It follows that $\bar{z}'S^k\bar{z} = \bar{z}'S^k\bar{x} = \bar{x}'S^k\bar{x} = \bar{z}'s^k$, and that $S^k\bar{z} = S^k\bar{x} \geqslant s^k$. Thus $\bar{z} \geqslant 0$ is also an optimal solution. If $\bar{r}_k' > 0$ then uniqueness implies $\bar{x} = \bar{z}$. Clearly in that case $(\bar{x},\ S^k\bar{x} - s^k)$ solves Q_k^4. If $(\bar{x},\ \bar{y})$ solves Q_k^4, then set $\bar{z} = \bar{x}$, and the KT-conditions are satisfied. Q.E.D.

The general problem of finding solutions (x,y) to a system $Tx - y = z$, in which the PSD square matrix T and the vector z are known is a familiar one. We use the terminology of Dantzig and Cottle (1967). A solution (x,y) is called complementary if $x_i y_i = 0$ for all i. The problem is important in different contexts: finding the solution to a system of linear inequalities with minimum norm (Ky Fan 1956, p 129), finding feasible directions which make a minimum angle with the gradient vector (Zoutendijk 1960, p 80-90), investigating symmetry and self-duality in QP (Dorn 1961, Cottle 1963). Dantzig and Cottle (1967) study the problem from a more general point of view and investigate the relations with LP, QP, and duality theory. The principal results are: if T is PD then the system has a non-negative complementary solution (Dorn), if T has all principal minors positive then the system has a non-negative complementary solution (Dantzig & Cottle), if T is PSD and the system has a non-negative solution then it also has a nonnegative complementary solution (Cottle).

Dantzig and Cottle also give an algorithm for finding non-negative complementary solutions, but this algorithm is just a special case of the ones proposed earlier by Zoutendijk. Consider the tableau

$$\begin{matrix} & \left[\quad x = 0 \quad\right] \\ \begin{bmatrix} y \\ = \\ -s^k \end{bmatrix} & \begin{bmatrix} \\ -S^k \\ \\ \end{bmatrix} \end{matrix}$$

We proceed just like in the dual simplex method, but we pivot only on the diagonal elements of the tableau. The results of Dantzig, Cottle, and Zoutendijk prove that it is always possible to find a suitable pivot element, and that the procedure is finite if we use an anti-degenracy precaution. Zoutendijk points out, however, that in practice cycling is very rare, and that in most cases it is better simply to choose the most negative element from the right-hand side in determining the main row. This 'complementary pivot' algorithm was also investigated by Dantzig (1963 p 490-497), and Lemke (1962) for solving the general QP-problem. Zangwill (1969 p 204-208) gives a geometric interpretation by relating it to optimization in a suitable manifold.

It is important which one of the problems Q_k^4 we solve first, i.e. how do we choose our first k. In general it seems good practice to search the matrix R for the most outlying edge. In particular we may start with Q_1^4, where 1 is the index corresponding with nI - E. An interesting problem is to find the maximin-r solution for the cone of all PSD-matrices. The edges of this cone are all PSD-matrices of rank one, and for C_1 = I we obtain $r(xx',I) = \frac{1}{n}\sqrt{n}$ for all x. Thus $r_{AT} = \frac{1}{n}\sqrt{n}$, and the AT-solution is I. For the cone of all SDC-PSD matrices we obtain in a similar way C_1 = nI - E, and $r_{AT} = \frac{1}{n-1}\sqrt{n-1}$. Thus nI - E is the center of the cone of all SDC-PSD matrices (in the AT-sense), but it is, at the same time, a very special edge of our cone of solutions (that is: almost always). The cone of DCT-matrices, a subcone of the cone of SDC-PSD matrices, has the same r_{AT} and AT-solution, which means geometrically something like: the cone DCT is symmetrically placed in the cone SDC-PSD.

D10

The next problem we have to solve is: suppose all edges are not known in advance, what can we do ? Of course this situation arises when the number of edges is not finite. Imagine two players A and B engaged in the following game: A is trying to find the AT-solution for a pointed cone P, but he does not know the cone. B knows what points belong to the cone. The two players are opposing each other: B is always trying to make A's current solution as bad as possible, or, equivalently, he is always trying to show that A's current solution is not the AT-solution. But at the same time he can only do this by revealing vital information about the cone, which makes it possible for A to improve his solution. The procedure is as follows

01) A produces a trial solution X^0, $k = 0$.

02) B finds the solution Y^k in the cone which minimizes $r(Y,X^k)$, where X^k is A's current solution.

03) A uses all Y^l $(l = 0,\dots,k)$ to compute the AT-solution, this is X^{k+1}, $k = k+1$, go to step 02.

If $(r_{AT}, C_1, \mathcal{S}')$ solves AT, then $X^{\infty} = C_1$, $Y^{\infty} \in \mathcal{S}'$. Moreover there are two sequences of r-values. In the second step we find

$$r^0_{min} \leq r^1_{min} \leq \dots \leq r_{AT},$$

and i the third step

$$r^0_{AT} \geq r^1_{AT} \geq \dots \geq r_{AT}.$$

In each iteration of the procedure (each pair of moves) we find (converging) upper and lower bounds for r_{AT}, if they are close enough we may stop.

Observe that we did not require that the cone must be polyhedral. If it is then we have, of course, finite convergence. If it is not then there will be an index l such that $\mathcal{S}' \subset \{Y^0, Y^1, \dots, Y^l\}$, and convergence is finite too. In this case A still does not know the cone P, but the AT-solution for P is identical to the one for the cone spanned by the Y^l. B cannot find an edge of P with $r^{l+1}_{min} < r^l_{AT} = r_{AT}$, his additional information does not help hin anymore. On the other hand if A should try out a subset with $r^{l+1}_{AT} > r^l_{AT} = r_{AT}$, then B is sure to come up with with $r^{l+2}_{min} < r^{l+1}_{AT}$. The system is stable, the game has ended in a draw. This is a consequence of the fact that a mixed version of the minimax theorem applies, the sequences A and B converge to the same value, it does not matter who makes the first move.

This is a nice method. Solving max-min problems can be fun. Alas, most of the fun disappears if we consider the amount of computation involved. The problems A has to solve in each move do not seem too prohibitive, but poor B has to minimize a pseudo-concave function on a nonpolyhedral convex set in each move. There are some slightly comforting circumstances. The function $r(Y,X^k)$ does not deviate much from linearity for fixed X^k: if the situation is not too degenerate then $Tr(Y^2)$ will be approximately equal for all vertices. Minimizing the numerator will bring us a long way. If $r^k_{min} > 0$ then we may also maximize the denominator and add the condition that the numerator is some positive constant. This is a QP-maximizing problem with a convex objective function. The maximum will be on an extreme point, and the algorithm discussed in appendix B applies.

In our ERMS-problem the equal-Δ method can not be applied. This is simply because not all edges can be transformed to Goode-shape, as in the polyhedral case. A slight modification makes the method more general and theoretically more sound. In the polyhedral case we transform the edges to unit lenght, take their average, and make this average vector of unit lenght too. Suppose the unit lenght edges are $S_1,\ldots,S_p$. We solve

$$\sum_{i=1}^{p} Tr(XS_i) \quad \max !$$

$$Tr(X^2) = 1.$$

The solution is proportional to $\sum S_i$. More generally: suppose the cone is pointed, and solve

$$\left[\sum Tr(XS_i)^q\right]^{1/q} \quad \max !$$

$$Tr(X^2) = 1.$$

For q = 1 we find our 'centroid' solution (that is the centroid of the edges, not of the cone !), for q = 2 we obtain $X = \sum \lambda_i S_i$, where the $\lambda_i > 0$ are the elements of the first eigenvector of $R = \{r_{ij}\} = \{Tr(S_i S_j)\}$, for $q \to \infty$ X can be taken as anyone of the edges (compare the complete method), for $q \to -\infty$ the solution X converges to the AT-solution. For all q the problem can be formulated as a game similar to the one we discussed previously. B's task is the same for all q, A's task varies and it extremely simple for q = 1 or q = 2.

For our one-dimensional example the maximin-r solutions are easy to find. We have only two edges, and X_{AT} is the solution for which $\rho^2(X_1, X_{AT}) = \rho^2(X_2, X_{AT})$. Evidently X_{AT} is also the centroid solution. For the edges X_1 and X_2 (page 60) we obtain

$$X_{AT} = (1 + \sqrt{3} \quad -2 - \sqrt{3} \qquad 1),$$
$$r_{AT} = \rho^2_{AT} = .933.$$

For the edges X_1 and X_4 (page 62)

$$X_{AT} = (4 + \sqrt{7} \quad -5 - 2\sqrt{7} \qquad 1 + \sqrt{7}),$$
$$r_{AT} = \rho^2_{AT} = .972.$$

Thus: no matter what other solution y we consider, $\rho^2(y, X_{AT})$ is always at least .972.

In the example on page 32-34 the AT-solution is easy to find by inspection. The following simple result is helpful.

<u>Theorem 4.6</u>: If Q_k is a pointed cone containing P with edges $T = \{T_1, \ldots, T_p\}$, if X_k is the AT-solution for Q_k, if T' is the subset of T for which $r(X_k, T_i) = r^k_{AT}$, and if all elements of T' are PSD, then X_k is the AT-solution for P and $r_{AT} = r^k_{AT}$.

Proof: If the edges of Q_k are PSD, then they are by definition edges of P. It follows from theorem 4.2 that $X_k \in P$. If $Y \in P$ then $Y \in Q_k$ and $r(Y, X_k) \geq r^k_{AT}$. In particular this is true for all edges of P not in T'. By theorem 4.2 again X_k is the AT-solution for P. Q.E.D.

Observe in the first place that (6), at the bottom of page 33, is not an edge of P. It is not even an edge of Q_2. This follows from the fact that (6) = (3) + 2 x (5), it also follows from corollary 3.6. Compute the AT-solution for Q_2. We try the average of (4) and (5) as a first guess. The r-values with the edges are

(1)	(2)	(3)	(4)	(5)
.89	.88	.88	.79	.79

It follows that this average is the AT-solution for Q_2, and because (4) and (5) are PSD it is also the AT-solution for P. The configuration is shown in figure 6a. For the corresponding one-dimensional problem the inequalities define two cones: one with edges

$$(2 \ -1 \ -1) \qquad\qquad (1 \ -1 \ \ 0),$$

and the other one with edges

$$(2 \ -1 \ -1) \qquad\qquad (1 \ \ 0 \ -1).$$

The union of these cones is the cone spanned by

$$(1 \ -1 \ \ 0) \qquad\qquad (1 \ \ 0 \ -1).$$

The AT-solution is the average of these two, $\rho^2_{AT} = .75$, and the configuration is shown in 6b.

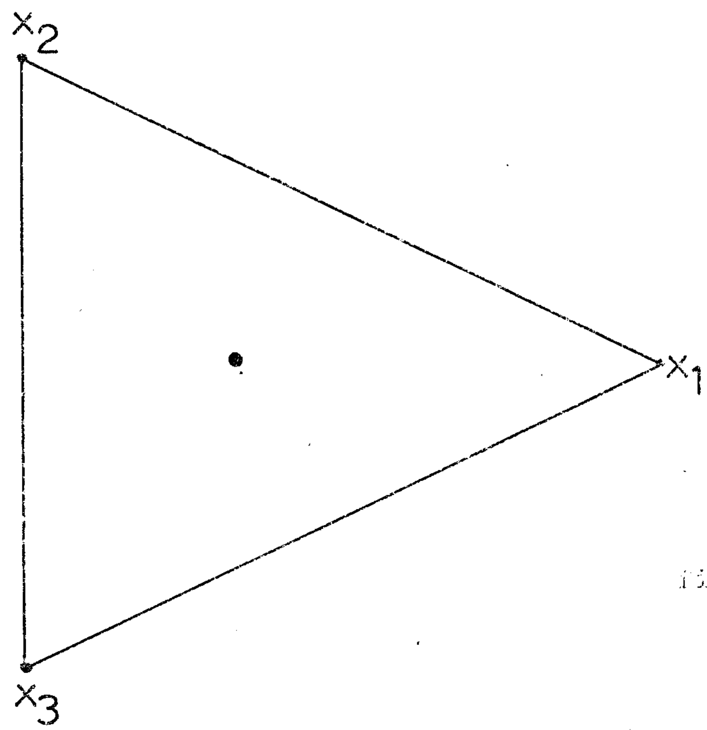


fig 6a

fig 6b

5 A comparison of different algorithms

The algorithms that are used most in NMS and related areas are projection algorithms. We give a general formulation that applies to both metric and nonmetric scaling theories, using the notation and terminology of De Leeuw (1970).

Some of the fundamental elements of each scaling theory are a mapping of the possible data structures into the algorithmic space

$$\tau : \Delta \rightarrow \mathfrak{S},$$

and a mapping of the possible representations into that same space

$$\phi : \Omega \rightarrow \mathfrak{S}.$$

Projection algorithms need angle-structures and least squares notions. Therefor we assume that $\mathfrak{S}$ is a Hilbert space. The loss function λ_δ is defined as

$$\lambda_\delta(\omega) = \inf_{s \in \tau(\delta)} \|\phi(\omega) - s\|,$$

and the scaling problem is

$$\inf_{\omega \in \Omega} \inf_{s \in \tau(\delta)} \|\phi(\omega) - s\|.$$

In nonmetric problems the set $\tau(\delta)$ usually is a convex cone, in metric problems it will, in most instances, be a single point. The structure of the set $\phi(\Omega)$ can be quite complex. In the Kruskal-Roskam-Shepard (KRS) algorithms $\mathfrak{S}$ is the real linear space in which each of the $\binom{n}{2}$ distances defines an axis, and Ω is the set of all n x p configuration matrices with centroid in the origin and $\|\omega\|^2 = 1/2n$ (or $\sum\sum d_{ij}^2 = 1$). The set $\tau(\delta)$ is a polyhedral convex cone in $\mathfrak{S}$, $\phi(\Omega)$ is a subset of the positive orthant of $\mathfrak{S}$. If ω_1 and ω_2 are related by a similarity transformation, then $\phi(\omega_1) = \phi(\omega_2)$, and consequently $\lambda_\delta(\omega_1) = \lambda_\delta(\omega_2)$. The set of perfect solutions is defined as $\phi^+(\tau(\delta) \cap \phi(\Omega))$, where ϕ^+ is the upper inverse of ϕ (the intersection may be empty). The algorithmic problem is to find a representation ω and a vector $s \in \tau(\delta)$ such that $\lambda_\delta(\omega)$ is as small as possible. A simple geometrical interpretation is available: $\lambda_\delta(\omega)$ is the sine of the angle between ω and s, consequently our problem can also be formulated as: find the vectors ω and δ in such a way that the angle between them is minimal.

My NMSPOM algorithms are conceptually almost identical to KRS. There is only one important difference: we work in a different algorithmic space, each inequality defines a dimension, and $\tilde{\tau}(\delta)$ is the positive orthant of the space (Guttman's absolute value principle also uses this space). We shall not discuss the LS version of NMSPOM in this chapter, because it has most of the technical advantages and disadvantages of the KRS-approach. It is undoubtedly more artificial and it should be applied only to some very special cases.

The GL-SSA series is similar to the KRS-series in most respects, but the loss-function is defined differently. The algorithmic space, the mappings τ and ϕ, the sets Δ and Ω are identical. Denote $\phi(\omega)$ by d, let $\mathcal{P}$ be the set of all permutation matrices of order $\binom{n}{2}$, and let

$$D^{*} = \{ x \mid (\exists_{\mathcal{P}} P) : x = Pd \}.$$

Write P for the cone $\tau(\delta)$, and let $P^{*} = P \cap D^{*}$. Evidently P^{*} is not empty. We now define

$$\lambda_{G}(\omega) = \inf_{d^{*} \in P^{*}} \| d - d^{*} \| .$$

This is Guttman's rank image principle. The mathematical and numerical properties of the algorithms based (exclusively) on it are quite bad, the theoretical foundation is much weaker than that of KRS. It should only be used in the initial iterations of mixed algorithms in the sense of Zangwill (1969 p 126-129), the 'basic algorithmic map' in projection type programs must always be KRS. Of course there is no objection at all to use GL a finite number of times in the iterations (and this is exactly what is done in Roskam's MINISSA-series).

Another interesting proposal is due to Sydow (1968). The principle is simple. Select a $\bar{d} \in \tau(\delta)$ and define

$$\lambda_{S}(\omega) = \| d - \bar{d} \|.$$

If we have numerical dissimilarities δ_{ij} we may choose $\bar{d} = \delta$ (LS-approach in KRS-space), if we have only order relations we may select the AT-solution for the cone $\tau(\delta)$. This is Sydow's proposal. Because the AT-solution is only determined up to interval scale level, Sydow defines

$$\lambda_{\delta}(\omega) = \inf_{\beta \in Re} \| d - \bar{d} + \beta e \| .$$

The positive orthant method was characterized by

$$\lambda_{\delta}(\omega) = \inf_{s \in P} \| t - s \|,$$

where P is the positive orthant of the inequality space and $t = \phi(\omega)$. A slight generalization gives

$$\lambda_{\delta}(\omega) = \inf_{s \in P} \| t - s \|_q ,$$

where $\|x\|_q$ is the l^q-norm ($q > 0$). It is easy to prove that for all q the minimum is attained for

$$s = \tfrac{1}{2}(t + |t|),$$

and consequently

$$\lambda_{\delta}(\omega) = \tfrac{1}{2} \| t - |t| \|_q .$$

We define Ω as the set of all n x p configuration matrices with centroid in the origin and $\frac{1}{2}\| \phi(\omega) \|_q = 1$. What interests us here is the limiting behaviour of the loss function for $q \rightarrow 0$. In De Leeuw (1970a) it is proved that the algorithm for $q \rightarrow 0$ becomes equivalent to the following one. Define

$$\lambda_{\delta}(\omega) = \frac{n_-}{n_- + n_+},$$

where n_- (n_+) is the number of negative (positive) elements in $\phi(\omega)$. We are looking for the maximum solvable subset, or: we maximize a variant of Kendall's tau. De Leeuw (1968) and Spence (1969) tried to do just this by direct maximalization and failed.

In this paper we have considered three main algorithms. The first one is NMSEMS (chapter 1), the second one the LS-method (section 4.2), and the third one AT (section 4.4). All these algorithms have an important property in common. They work in the algorithmic space of SDC-matrices. What happens if we translate the KRS-rationale into that space ? Let Q_0 denote the cone of SDC matrices obtained by the method of section 3.7. Then we define

$$\lambda_{\delta}(X) = \inf_{C \in Q_0} \| XX' - C \| ,$$

and the algorithmic problem is (if χ_p is the set of all real centered n x p with orthogonal columns and $Tr(X'X)$ equal to some constant c).

$$\inf_{X \in \chi_p} \lambda_{\delta}(X) = \inf_{X \in \chi_p} \inf_{C \in Q_0} \| XX' - C \| .$$

This suggests the following algorithm. Let $\overline{Q}_0$ be the compact convex polyhedron defined by $\overline{Q}_0 = Q_0 \cap \{ C \mid Tr(C) = c \}$.

01) Start with some $\overline{C}^1 \in \overline{Q}_0$, k = 1.

02) Solve

$$\inf_{X \in \chi_p} \| XX' - \overline{C}^k \| .$$

This is equivalent to minimizing

$$Tr((XX' - \overline{C}^k)'(XX' - \overline{C}^k)).$$

Necessary for an extreme value is

$$\overline{C}^k X = X(X'X),$$

and if we assume that $\overline{C}^k$ has at least p nonnegative eigenvalues, then necessary and sufficient for a minimum (the global minimum) is $X^k = K\Lambda^{\frac{1}{2}}$, in which Λ and K contain the p largest eigenvectors and their eigenvalues.

03) Compute $\tilde{C}^k = X^k(X^k)'$.

04) Solve

$$\inf_{C \in \overline{Q}_0} \| C - \tilde{C}^k \| .$$

This is a QP-problem. If the vertices of $\overline{Q}_0$, T_p, are known, then the problem is equivalent to

$$\lambda' R \lambda - 2 \lambda' r^k + \beta \text{ min! } \lambda \geq 0,$$

with $r_{pq} = Tr(T_p T_q)$, $r_p^k = Tr(T_p \tilde{C}^k)$, $\beta = Tr[(\tilde{C}^k)^2]$. The KT conditions are: there are multipliers $\mu \geq 0$ such that

$$\mu = R\lambda - r^k,$$

$$\mu'\lambda = 0, \mu \geq 0, \lambda \geq 0.$$

which suggest applying the complementary pivot algorithm to the initial tableau

$$\begin{bmatrix} \mu = -r^k \end{bmatrix} \quad \begin{bmatrix} \lambda = 0 \\ -R \end{bmatrix}.$$

Of course R is PSD. The solution is $\overline{C}^{k+1}$, $k = k + 1$, go to step 02.

Under the assumption that $\overline{C}^k$ has p nonnegative eigenvalues for each k, this algorithm converges to at least a local minimum. The assumption is not too restrictive, p will usually be small compared to n. In this form it is also very easy to decide whether or not we include nI − E among the T^p. This is a trasnlation of the KRS-rationale into C-space, the algorithm is the translation of the semi-nonmetric phase of Young's TORSCA (or equivalently De Leeuw's ALS) into this space. Observe that in KRS-space these methods do not converge, and consequently they too must be used only a finite number of times in a convergent KRS-program (the TORSCA-programs do this).

There are other ways to formulate the problem. If we use

$$\inf_{X \in \mathcal{X}_p} \| XX' - \overline{C}^k \| = \sqrt{\sum_{i=p+1}^{n} \lambda_i^2},$$

where the λ_i are the n − p smallest eigenvalues of $\overline{C}^k$, then we can write

$$\inf_{X \in \mathcal{X}_p} \inf_{C \in \overline{Q}_0} \| XX' - C \| = \inf_{C \in \overline{Q}_0} \sqrt{\sum \lambda_i^2}.$$

Unfortunately the root-mean-square of the n − p smallest eigenvalues is not a quasiconcave function. Otherwise the vertex T^p with the smallest function value would be the solution, computations would be extremely simple, local minima would not be possible. An interesting question, which can only be answered numerically, is: how well does the best vertex perform in terms of Kruskal's stress ? A very simple algorithm, which only needs the vertices T^p and a fast subroutine to compute the eigenvalues (not the eigenvectors !), is: start with the best vertex as current solution, then use a cyclic method by searching on the line segment connecting the current solution with all vertices in turn for better solutions. Again we should like to know how well the best solution performs in terms of stress.

The Sydow-method, when translated into SDC-space, simplifies considerably. We must solve

$$\inf_{X \in \mathcal{X}_p} \| XX' - C^0_{AT} \| ,$$

where C^0_{AT} is the AT-solution for C_0. The solution consists simply of the first p eigenvectors of C^0_{AT}, suitably scaled.

In NMSMS we have the comparable problem

$$\inf_{X \in \mathcal{X}_p} \| XX' - A \| ,$$

only there is no guarantee that $A \in C_0$. Geometrically NMSMS has no satisfactory rationale, its merits must be investigated numerically. Some numerical results have been reported by De Leeuw (1968). Carroll and Chang have compared the performance of a maximum sum algorithm and a positive orthant algorithm for the NMFA-problem, and found (to their surprise) that MS did a better job.

This is related to the principal disadvantage of projection-type methods. If perfect solution in p dimensions do not exist then requiring weak order isometricity and measuring the errors has the effect that the configurations become partially degenerate. Tied distances are produced at places where strong order isometricity would forbid them. Letting $q \rightarrow 0$ in POM is an attempt to overcome this disadvantage, but it is costly and as long as $q > 0$ degenerate solutions remain possible. The rank image principle is another heroic attempt to preserve strong monotonicity, but it fails too. In NLSMMS degeneracy does not seem to be a serious problem. In fact cases occur in which the inherent degeneracy in the data can be completely removed from the solution simply by deflating A (remove the first principal component).

If perfect solutions in p dimensions are available (this is the exception rather than the rule with the usual choice of p) then the projection-type methods will produce a completely arbitrary perfect solution (in the weak sense), and we do not know how representative this solution is. This disadvantage becomes important if there are relatively few constraints (as in

USA-II), and the solution set is large. In this case it seems wise to collect some information about the uniqueness (by using a number of different random starting points).

The methods in this paper concentrate on the C-matrix, not on the configuration. The solution is a particular PSD-SDC matrix, from which the configuration can be obtained by principal components methods. This involves a quite different emphasis. It implies that a perfect solution always exists (in most cases of interest), and that the main point is finding a representative solution, which turns out to be strongly order isometric in most cases.

I finish with some practical advice. In any NMS problem one should start with the EMS solution. This solution can be used for three purposes: in the first place we get an idea about the dimensionality. In the second place we get an initial configuration for KRS or PCH. And in the third place the EMS result is itself an interesting NMS-solution. The next candidate for standard application is a projection algorithm, with as the most advanced species a program from Roskam's MINISSA-series. In choosing the options we must remember never to use a random initial configuration, we must never use rank images in our basic mapping, and we must never use stress formula one. Alternatively we can use the analogue in C-space developed in this section. We can get a quick idea about uniqueness by computing the AT-solution for C_0 (Sydow's method in C-space) and the AT-solution for the cone defined by the T-method of section 3.2. The complete AT-method cannot be recommended, because it is highly impractical. Wait for the next generation of computers. The LS-method of section 4.2 seems worthwhile if we have numerical dissimilarities.

Appendix A: Additional definitions

This appendix gives some additional concepts which properly belong in the section on preliminaries. The definition of a PCC-distance is changed. We start with an arbitrary set A (finite or infinite), and a set $B \subset A \times A$.

Definition A1: A function $\theta : B \to Re$ is called a quasi-metric iff θ has a real valued extension $\bar{\theta}$ to $A \times A$ such that

A1.1: $\bar{\theta}(A \times A)$ has a least element λ.

A1.2: $(\forall_A a)(\forall_A b)$: $\bar{\theta}(a,b) = \bar{\theta}(b,a)$.

A1.3: $(\forall_A a)(\forall_A b)$: $\bar{\theta}(a,b) = \lambda \Leftrightarrow (\forall_A c)$: $\bar{\theta}(a,c) = \bar{\theta}(b,c)$.

If there is an extension satisfying A1.1-A1.3 and in addition

A1.4: $(\forall_A a)(\forall_A b)(\forall_A c)$: $\bar{\theta}(b,a) + \bar{\theta}(b,c) \geq \bar{\theta}(a,c) + \bar{\theta}(b,b)$,

then θ is called a semi-metric. If there is an extension satisfying A1.1-A1.4 and

A1.5: $\lambda = 0$,

then θ is called a metric. If for all extensions $\bar{\theta}$ it is true that

A1.6: $\bar{\theta}(A \times A)$ has no more then two elements,

then θ is called degenerate.

Two formula's which can be derived from A1 are

01) $(\forall_A a)(\forall_A b)$: $2\,\bar{\theta}(b,a) \geq \bar{\theta}(a,a) + \bar{\theta}(b,b)$,

02) $(\forall_A a)$: $\bar{\theta}(a,a) = \lambda$.

We can construct the implication diagram

$$\begin{tikzcd} A1.4 \arrow[r] \arrow[dr] & A1.2 & \\ & 01 \arrow[r] & A1.1 \\ (\exists_{Re}\lambda):\ A1.3 \arrow[r] & 02 \arrow[ur] & \end{tikzcd}$$

We have used the concept of an extension of a real valued function simply because there may be missing data. Of course we can consider the subset B as a binary relation on A. This gives us the following definition.

Definition A2: A function defined on a subset B of a cartesian product $A \times A$ is called

A2.1: complete iff B is full (i.e. $B = A \times A$), A2.2: <u>inclusively triangular</u> iff B is antisymmetric and reflexive,

A2.3: <u>exclusively triangular</u> iff B is asymmetric and irreflexive,

A2.4: <u>rectangular</u> iff $B \subseteq A_1 \times A_2$, where $\{A_1, A_2\}$ is a partitioning of A.

An easy consequence of these definitions is that all exclusively triangular and rectangular real valued functions are automatically quasi-metrics (at least if A is finite). If A is finite and B is irreflexive then δ is not a quasi-metric iff $(\exists_A a)(\exists_A b)$: $\delta(a,b) \neq \delta(b,a)$.

The next primitive is a mapping ϕ of a subset D of $A \times A$ into the poset $\langle \Delta, \geq_o \rangle$.

<u>Definition A3</u>: The system $\langle D, \Delta, \phi, \geq_o \rangle$ is called a <u>partially ordered generalised distance</u> (or: <u>POG-distance</u>) iff there exists a quasimetric $\delta: D \to Re$ such that

A3.1: $(\forall_D(a,b))(\forall_D(c,d))$: $\phi(a,b) >_o \phi(c,d) \Rightarrow \delta(a,b) > \delta(c,d)$,

A3.2: $(\forall_D(a,b))(\forall_D(c,d))$: $\phi(a,b) =_o \phi(c,d) \Rightarrow \delta(a,b) = \delta(c,d)$.

If A3.1 and A3.2 can be replaced by the stronger condition

A3.3: $(\forall_D(a,b))(\forall_D(c,d))$: $\phi(a,b) \geq_o \phi(c,d) \Leftrightarrow \delta(a,b) \geq \delta(c,d)$,

then the system is called a <u>weakly ordered generalised distance</u> (or: <u>WOG-distance</u>). A POG-distance is <u>degenerate</u> iff all quasi-metrics satisfying A3.1 and A3.2 are degenerate.

Again it follows that if A is finite then all systems in which ϕ is exclusively triangular or rectangular are POG-distances.

We want to compare our new definition of a POG-distance with the one on page 1. In the first place our previous definition was only meant for the case of finite A. In the second place part 0.1.3 is not quite correct, because it assumes that all δ_{ij} and δ_{kl} can be compared. We restate the three parts of 0.1 as follows: there exists an extension $\phi_o : A \times A \to \Delta$ and a weak order extension $\geq_1$ of $\geq_o$ such that

$$01: (\forall_A a)(\forall_A b): \phi_e(a,b) \geqslant_1 \phi_e(b,a) \Leftrightarrow \phi_e(b,a) \geqslant_1 \phi_e(a,b).$$

$$02: (\forall_A a)(\forall_A b)(\forall_A c): \neg\,(\phi_e(a,a) \succ_1 \phi_e(b,c)).$$

$$03: (\forall_A a)(\forall_A b)(\forall_A c): \phi_e(a,a) \geqslant_1 \phi_e(b,c) \Leftrightarrow (\forall_A d): \phi_e(b,d) =_1 \phi_e(c,d).$$

Of course 01 is equivalent with

$$01: (\forall_A a)(\forall_A b): \phi_e(a,b) =_1 \phi_e(b,a),$$

and 02 is equivalent to

$$02: (\forall_A a)(\forall_A b)(\forall_A c): \phi_e(a,a) \leqslant_1 \phi_e(b,c).$$

Moreover 02 implies that $\phi_e(A \times A)$ has a least element λ. Take $x, y \in A$ and substitute $x = a$, $b = c = y$ in 02. Then $\phi_e(x,x) \leqslant_1 \phi_e(y,y)$. If $y = a$, $b = c = x$, then $\phi_e(y,y) \leqslant_1 \phi_e(x,x)$. Thus 02 also implies that

$$(\forall_A a)(\forall_A b): \phi_e(a,a) =_1 \phi_e(b,b).$$

The reverse implication in 03 (and the similar one in A.1.3) is not essential, it simply must be interpreted as a definition of equality in A:

$$a =_2 b \text{ iff } (\forall_A c): \phi_e(c,a) =_1 \phi_e(c,b).$$

Obviously $=_2$ is an equivalence relation, and $\phi_e(a,b) =_1 \lambda$ iff $a =_2 b$. We have proved that ϕ_e has all the properties of a complete quasimetric, except for the fact that it is not necessarily real-valued. In the infinite case, with connected $\geqslant_1$, order-isomorphic imbedding into the reals is possible iff Δ is seperable in its order topology iff Δ contains a countable order-dense subset (Cantor-Birkhoff-Debreu theorem). In the finite case the system is a POG-distance in the sense of definition A3 iff 01, 02, 03 are true.

Definition 0.1 must be abandoned. Consider the conditional case

$$\delta_{11} \leqslant_c \delta_{12} \leqslant_c \delta_{13}$$
$$\delta_{22} \leqslant_c \delta_{23} \leqslant_c \delta_{21}$$
$$\delta_{33} \leqslant_c \delta_{31} \leqslant_c \delta_{32}$$

Requirements 0.1.1, 0.1.2, and 0.1.3 are satisfied. But the quasimetric would have to satisfy the inequalities

$$\left.\begin{array}{l} \theta_{12} < \theta_{13} \\ \left.\begin{array}{r} \theta_{23} < \theta_{21} = \theta_{12} \\ \theta_{31} = \theta_{13} < \theta_{32} = \theta_{23} \end{array}\right\} \Rightarrow \theta_{13} < \theta_{12} \end{array}\right\} \Rightarrow \text{contradiction.}$$

Although O1, O2, and O3 would be satisfied if we simply use $\succsim$, an extension $\succsim$ as required by O1, O2, O3 does not exist.

We need another definition that is useful in proving representation theorems.

Definition A4: If $A_s = \{a_1, a_2, \ldots, a_n\}$ is a subset of A and δ is a real-valued function on A x A then the matrix

$$t_{ij} = -\tfrac{1}{2}\left\{ \delta^2(a_i,a_j) - \frac{1}{n}\sum_l \delta^2(a_i,a_l) - \frac{1}{n}\sum_k \delta^2(a_k,a_j) + \frac{1}{n^2}\sum_k\sum_l \delta^2(a_k,a_l)\right\}$$

is called the pseudo-scalar-product (or PSP) matrix associated with A_s.

It is easy to see that if δ is a complete quasimetric then T is SDC and consequently of rank $\leq n - 1$. Another useful formula is

$$t_{ii} + t_{jj} - 2t_{ij} = \delta^2(a_i,a_j).$$

Next we give a definition used in metric multidimensional scaling methods.

Definition A5: If δ is a real-valued function of $B \subset A \times A$, then a representation $\chi(A)$ is said to be linearly isometric iff there is an extension $\bar{\delta}$ of δ to A x A and a real number ρ such that

$$(\forall_A a)(\forall_A b): \bar{\delta}(a,b) + \rho = d(\chi(a),\chi(b)).$$

It is called quadratically isometric iff

$$(\forall_A a)(\forall_A b): \bar{\delta}^2(a,b) + \rho = d^2(\chi(a),\chi(b)),$$

and it is said to be identically isometric (or congruent) iff

$$(\forall_A a)(\forall_A b): \bar{\delta}(a,b) = d(\chi(a),\chi(b)).$$

Such representations in p dimensions are called, respectively, linear, quadratic, or congruent p-representations.

Definition A6: A p-representation $\chi(A)$ is called irreducible iff there is no configuration in (p-1)-space with the same set of distances.

Appendix B: A summary of representation theorems

D1.6

While writing this appendix the paper Lingoes (1969) has come to my attention. It contains a proof of Guttman's (n-2)-theorem. Although the method of proof differs from our method, and our results are more simple, more elegant, and much more general, there are obvious similarities. One of the main differences is that we start with the theorem for the congruent case, then investigate the linear and quadratic cases, and the theorems covering the nonmetric case come out as corollaries.

The euclidean representation problem as two different aspects. In the imbedding problem (metric version) we want to find necessary and sufficient conditions that an arbitrary quasi-metric space must satisfy in order to be congruent to a subspace of euclidean p-dimensional space E_p. The space problem, which is less general, is simply to give conditions that a quasi-metric space must satisfy in order to be congruent to E_p as a whole. More precisely: the imbedding problem is to find a representation $\chi(A)$ in E_p such that

$$(\forall_A a)(\forall_A b): \quad \theta(a,b) = d(\chi(a),\chi(b)).$$

In the space problem we want this to be the case too, but in addition we require

$$(\forall_{E_p} x)(\forall_{E_p} y): \; d(x,y) = \theta(\chi^{-1}(a),\chi^{-1}(b)).$$

In the space problem the function χ must be continuous and one-to-one, with a continuous and one-to-one inverse. Consequently it must be a homeomorphism. If we have solved the imbedding problem then all we have to do to solve the space problem is to characterize E_p among its subsets (cf Blumenthal 1953 p 90-91). In the quasi-metric case both problems have been completely solved for the most important spaces: euclidean, real Hilbert, elliptical, generalized hyperbolic, and sperical space. The recent developments in the social sciences have led to a generalization of these problems to the POC-distance case.

In this paper we are mainly interested in the imbedding problem for finite A. This is a typically psychometric problem, which is relevant for the MDS scaling theory. In De Leeuw (1970b) it is argued that any scaling theory is incomplete without a complete solution to the finite case of the relevant imbedding problem. The space problem, however, is a problem which properly belongs in the area of the mathematical social sciences. The completeness of a scaling theory is independent of the fact whether the corresponding space problem (in case of a homeomorphism we may speak of the topological problem, cf the work of Debreu, Chipman, and Pfanzagl) is solved. The hypothesis that the cognitive space for a given infinite universe of discourse is congruent to E_p (or the weaker hypotheses that the two spaces are homeomorphic) is a scientific hypotheses, which is mathematically equivalent to the conditions described in the theorem that solves the space problem. The hypotheses can only be tested through its consequences, and one of the consequences is that all finite samples from the universe can be congruently imbedded in E_p. Of course our scientific hypotheses could also have been that a given infinite universe of discourse is congruent with a metric subspace of E_p (not a linear subspace, because this would be equivalent to the space hypotheses for E_r, $r < p$).

We start our list of representation theorems with the simplest case in which A is finite and a complete semi-metric $\mathcal{C}$ is given. Unfortunately the terminology in definition A1 is still not beyond reproach. It is better to interchange A.1.4 and A.1.5. A complete semimetric is then defined by

$$P1: (\forall_A a)(\forall_A b): \mathcal{C}(a,b) = 0 \Leftrightarrow (\forall_A c): \mathcal{C}(a,c) = \mathcal{C}(b,c).$$

$$P2: (\forall_A a)(\forall_A b): \mathcal{C}(a,b) = \mathcal{C}(b,a).$$

$$P3: (\forall_A a)(\forall_A b): \mathcal{C}(a,b) \geq 0.$$

The first representation theorem is due in all essentials to Schoenberg (1935). In the psychometric literature it is often attributed to Young and Householder (1938). Other necessary and sufficient conditions for the finite case were already found by Menger (1928), using determinants. Our formulation of the theorem is more consistent with the rest of the paper.

<u>Theorem B1</u>: If $\mathcal{C}$ is a real-valued function on A x A then an irreducibly congruent

p-representation exists iff

B.1.1: θ is a complete semi-metric.

B.1.2: The PSP-matrix of A is PSD of rank p.

Proof: The proof is simple, we only give an outline. Suppose there is a representation $\chi(A)$ in p dimensions such that $\theta(a_i, a_j)$ equals the euclidean distance between x_i and x_j. Some easy algebra gives $t_{ij} = (x_{i.} - \bar{x})'(x_{j.} - \bar{x})$, with $\bar{x} = \frac{1}{n}\sum x_{i.}$. Thus T is PSD of rank p. If T is PSD of rank p then we can find a representation by decomposing T as T = XX' (with X n x p). In that case $d_{ij}^2 = \sum x_{is}^2 + \sum x_{js}^2 - 2\sum x_{is}x_{js} = t_{ii} + t_{jj} - 2t_{ij} = \theta^2(a_i, a_j)$. Q.E.D.

A solution of the embedding theorem for all A (finite or infinite) can be derived from the following result. It was proved by Menger (1928) and in a more refined form by Blumenthal (1953 p 35): A semimetric space (A, θ) is congruently imbeddable in E_p iff each subset of A containing p + 3 points is. The space problem was also solved by Menger: a semimetric space is congruent with E_p iff it is topologically complete, metrically convex, externally convex, and irreducibly congruently imbeddable in E_p. For the meaning of these terms we refer the reader to Blumenthal's books, where further refinements can be found.

Here we conclude the case in which θ is a complete semi-metric. The incomplete case can easily be solved using extensions of our quasi-metric. The development in psychometrics in the early fifties involved dissimilarities measured on an interval scale, and here the concepts of linear and quadratic isometricity become relevant. The theorems that must guarantee the existence of linear and quadratic p-representations are essentially easy substitutions into theorem B1. From these new theorems a few interesting corollaries follow. We may restrict ourselves to the case in which the relation B is irreflexive: if there exists a $c \in A$ such that $\theta(c,c)$ is defined, then a linear p-representation exists iff a congruent p-representation exists for $\tilde{\theta}(a,b) = \theta(a,b) - \theta(c,c)$.

<u>Theorem B2</u>: If A is a finite set of n elements, and θ is an irreflexive

real-valued function. Then a linear (n-1)-representation exists iff δ is a quasi-metric.

Proof: It is obvious that the condition is necessary. To prove sufficiency we assume that δ is a quasi-metric, choose a symmetric extension to all nondiagonal elements and define

$$e_{ij}^{\gamma} = \begin{cases} \bar{\delta}(a_i, a_j) + \gamma & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}$$

It is sufficient to prove that there is a γ such that the PSP matrix derived from E^{γ} is PSD. The PSP matrix V^{γ} can be written as

$$V^{\gamma} = A + 2\gamma B + \gamma^2 C$$

with both A, B, and C not dependent on γ and SDC. This representation is unique, and $C = (2n)^{-1}(nI - E)$. For the off-diagonal elements we have

$$v_{ij} = a_{ij} + 2\gamma b_{ij} - \gamma^2/2n.$$

It is easy to see that there is a value γ_{ij} such that $v_{ij} \leq 0$ for all $\gamma > \gamma_{ij}$. Let $\hat{\gamma} = \max_{i,j} \gamma_{ij}$. Let $\bar{\gamma} = -\min_{i,j} \bar{\delta}(a_i, a_j)$, and $\ddot{\gamma} = \max(\hat{\gamma}, \bar{\gamma})$. For $\gamma > \ddot{\gamma}$ the elements in E^{γ} are a complete semi-metric, and V^{γ} is an SDC T-matrix, and consequently PSD of rank $\leq n - 1$ (theorem 1.7). B2 now follows from B1. Q.E.D.

Observe that $nI - E$ again plays a crucial role. It is well known that a large additive constant gives us configurations which look like a regular simplex. In algebraical terms

$$\lim_{\gamma \to \infty} \frac{n(n-1)V^{\gamma}}{\mathrm{Tr}(V^{\gamma})} = nI - E.$$

Theorem B3: Suppose A is a finite set of n elements, and δ is an irreflexive real valued function, which is moreover non-negative. Then a quadratic (n-2)-representation exists iff δ is a nondegenerate quasi-metric.

Proof: We give a brief outline, using the notation of the previous proof. Now $V^{\gamma} = A + \gamma C$, with $A = C$ iff δ is degenerate. If A is PSD of rank $\leq n - 2$, we choose $\gamma = 0$. If A is PSD of rank $n - 1$, and $\bar{\lambda}$ is the smallest positive root, we choose $\gamma = -2\bar{\lambda}$. If A is indefinite, then let $\bar{\lambda}$ be the smallest negative root, and set $\gamma = -2\bar{\lambda}$. The resulting matrix V^{γ} is PSD of rank $\leq n - 2$. Apply theorem B1. Q.E.D.

Of course a quadratic p-representation is also a strong p-representation.

Corollary B4: A nondegenerate POG-distance always has a strong (n-2)-representation.

The next corollary gives an answer to the representation problem for linear-rank constraints (or general contrasts) and faithful solutions of arbitrary degree. Let $\mathcal{V} = \{V_1, V_2, \dots, V_l\}$ be a set of n x n matrices with $\mathrm{Tr}(EV_k) = 0$ for all $k = 1, \dots, l$.

Corollary B5: If there exists a quasimetric δ such that

$$\sum_i \sum_j v_{ij}^k \delta_{ij} > 0 \qquad k = 1, \dots, l$$

then there exists a representation $\chi(A)$ in n-1 dimensions such that

$$\sum_i \sum_j v_{ij}^k d_{ij} > 0 \qquad k = 1, \dots, l.$$

Corollary B6: If there exists a [nondegenerate] quasi-metric such that

$$\sum_i \sum_j v_{ij}^k \delta_{ij}^2 > 0 \qquad k = 1, \dots, l$$

then there exists a representation $\chi(A)$ in n-2 dimensions such that

$$\sum_i \sum_j v_{ij}^k d_{ij}^2 > 0 \qquad k = 1, \dots, l.$$

Observe that B4 follows from ~~both~~ B3 through B6, and observe that the converse of both B5 and B6 are also obviously valid.

It is not too easy to solve the imbedding problem in E_p ($p < n - 2$) for POG-distances. Of course we can say that the problem has an irreducible strong p solution iff there is a PSD matrix of rank p in the interior of the cone ψ_0, but this is hardly a satisfactory criterion. The space problem for POG-distances can be solved along the lines of Beals, Krantz, and Tversky. The main problem is to define metric betweenness in terms of the order on the dissimilarities. If this is done we can introduce a concatenation operation, and introduce axioms that guarantee that the distance set is a fully ordered, archimedean, cancellative, positive semigroup. This defines a ratio-scale metric, in terms of which the space problem can be easily solved. Because we are only interested in psychometric problems, we shall not give a precise analysis of the space problem.

Appendix C: Advanced computational remarks

The algorithm explained in sections 3.4 - 3.7 is cumbersome but sure to converge, the modifications discussed in 3.9 are faster but may fail to reach the optimal point. In this appendix we devise a different algorithm. The general formulation is probably new, but special cases were already applied by Murty (1968) and Cabot & Francis (1970). In fact the algorithm is nothing but a combination of the cutting-plane method with an extreme-point ranking procedure. The problem is

$$\min_{x \in S} f(x).$$

Call this problem P. We suppose that the following assumptions are met.

C1: S is a closed convex subset of Re^n.

C2: f is quasi-concave.

These two assumptions define the class of problems we are interested in. The next three assumptions are more technical and needed for algorithmic purposes.

C3: A closed convex polyhedron $A_0 \supset S$ is given.

C4: There is a linear functional $g(x)$ such that $g(x) \leq f(x)$ for all $x \in A_0$. Moreover $\min_{x \in A_0} g(x)$ exists and is finite.

C5: If x is a point not belonging to S we have a rule for constructing a hyperplane that separates x and S.

Some of the consequences of these assumptions are

01) $\min_{x \in S} f(x) \geq \min_{x \in A_0} f(x) \geq \min_{x \in A_0} g(x) > -\infty.$

02) $\min_{x \in A_0} g(x) \leq \min_{x \in S} g(x) \leq \min_{x \in S} f(x).$

03) If Y is the set of points for which f is minimal on A_0 then Y contains at least one of the extreme points of A_0. Three other true statements can be obtained by replacing f by g and/or A_0 by S in the previous sentence.

We are now ready to describe our algorithm. Start with problem Q_0:

$$\min_{x \in A_0} f(x).$$

To solve Q_0 we first solve Q_0^1:

$$\min_{x \in A_0} g(x),$$

which is an LP-problem, the solution x_0^1 is an extreme point of A_0. Compute $f(x_0^1)$. Evidently

$$\min_{x \in A_0} g(x) = g(x_0^1) \leq \min_{x \in A_0} f(x) \leq f(x_0^1).$$

If $g(x_0^1) = f(x_0^1)$ this evidently means that x_0^1 is the solution of problem Q_0. If this is not the case we find the extreme point of A_0 which gives the minimum value of $g(x)$ on all the extreme points of A_0, excluding x_0^1. A procedure to do just this is described by Murty (1968). This gives x_0^2. Again

$$\min_{x \in A_0} f(x) \leq f(x_0^2),$$

but not necessarily

$$g(x_0^2) \leq \min_{x \in A_0} f(x),$$

of course

$$g(x_0^1) \leq g(x_0^2).$$

Let $\underline{f}^2 = \min(f(x_0^1), f(x_0^2))$, and let f_{min}^0 and g_{min}^0 denote the minima of $f(x)$ and $g(x)$ over A_0. If $\underline{f}^2 < g(x_0^2)$, then

$$g(x_0^1) = g_{min}^0 \leq f_{min}^0 \leq \underline{f}^2 < g(x_0^2).$$

Suppose the minimum of $f(x)$ on A_0 occurs at the extreme point x_0^k. $(k>2)$ Then

$$f_{min}^0 = f(x_0^k) \leq \underline{f}^2 < g(x_0^2) \leq g(x_0^k),$$

and consequently $f(x_0^k) < g(x_0^k)$, which contradicts C4. Consequently either x_0^1 or x_0^2 minimizes $f(x)$ on A_0. A similar conclusion holds if $\underline{f}^2 = g(x_0^2)$. If $\underline{f}^2 > g(x_0^2)$ we find the next extreme point x_0^3, test if $\underline{f}^3 \leq g(x_0^3)$, and so on. Because the number of extreme points of A_0 is finite, there are only two possibilities. Either $\underline{f}^k \leq g(x_0^k)$ for some k, or at some stage all extreme points of A_0 have been investigated, and the solution is the point with the lowest f-value. The procedure ends in a finite number of steps. If Q_0 is solved, then we investigate if the solution $\bar{x}_0$ lies in S. If it does, we have solved P. If it does not, we make a cut by adding the constraint $h_0'x \leq c_0$. This defines a new polyhedron A_1, and we solve the problem Q_1. And so on.

Although the problems Q_s are solved in a finite number of steps, the solution $\bar{x}$ of P will in general not be found after solving a finite number of problems Q_s. We give a more formal statement of the algorithm.

i) Start with $k = 0$, $l = 1$, T_1 is the set of extreme points of A_k.

ii) Solve the LP-program $\min_{x \in T_l} g(x)$. This gives x_k^l, $g(x_k^l)$, and $f(x_k^l)$.

iii) Compute $\underline{f}_k^l = \min_{i=1}^{l} f(x_k^i)$.

iv) If $\underline{f}_k^l \leq g(x_k^l)$ then let $\bar{x}_k$ be any of the x_k^i $(i=1,\ldots,l)$ such that $f(x_k^i) = \underline{f}_k^l$. Go to step vi.

v) If $\underline{f}_k^l > g(x_k^l)$ then $T_{l+1} = T_l - \{x_k^l\}$. If $T_{l+1} = \emptyset$, then let $\bar{x}_k$ be any of the x_k^i $(i=1,\ldots,l)$ such that $f(x_k^i) = \underline{f}_k^l$, and go to step vi. If $T_{l+1} \neq \emptyset$, then $l = l + 1$, go to step ii.

vi) If $\bar{x}_k \in S$ then stop, $\bar{x}_k$ solves problem P.

vii) If $\bar{x}_k \notin S$, then $A_{k+1} = A_k \cap \{x \mid h_k'x \leq \alpha_k\}$, $k = k + 1$, go to step i.

The stopping criterium used is not satisfactory if we do not expect finite convergence. Then we should use (in step vi): if $\bar{x}_k$ is close enough to S, then stop, or: if k is larger then some maximal permitted value, then stop. For ease of reference we formulate the fundamental theorem again.

<u>Theorem C1</u>: Let $\underline{f}_k^l = \min_{i=1}^{l} f(x_k^i) = f(x_k^t)$, with t somewhere between 1 and l. Let $\bar{f}_k^l = g(x_k^l) = \max_{i=1}^{l} g(x_k^i)$. If $\underline{f}_k^l \leq \bar{f}_k^l$ then x_k^t solves problem Q_k.

We consider some special cases. If f(x) is linear and S is polyhedral, then we take $g(x) = f(x)$ and $A_0 = S$. The problem reduces to an LP-problem and the algorithm reduces to the particular LP-method used in step ii. If f(x) is linear, and $S = \{x \mid g_i(x) \leq a_i\}$ where the g_i are a number of convex differentiable functions, then our algorithm reduces to the convex cutting plane method or the supporting hyperplane method (depending on the rule mentioned in C5). If f(x) is quasi-concave, and $\max_{x \in A_0} g(x) < \min_{x \in A_0} f(x)$ then our algorithm requires an enumeration of all vertices, as the algorithm in section 3.4 does. The function $g(x)$ only serves as a tool for ranking the vertices. Other important special cases are bilinear programming (Mangasarian 1964, Mangasarian & Stone 1964), quasi-monotone programming (Martos 1965), concave quadratic programming (Cabot & Francis 1970), linear fractional programming (Martos 1965), and solving the fixed charge problem (Murty 1968).

In some of these problems the quasi-concave function $f(x)$ can be written as $f(x) = G(\sum x_i h_i(x))$, with G concave and homogeneous. If nonnegativity of the x_i is among our requirements, then we can easily find our functionals $g(x)$: let $u_i = \min_{x \in B} G(h_i(x))$, and $g_B(x) = \sum u_i x_i$, then $g_B(x) \leq f(x)$ for all $x \in B$.

In the paper of Cabot & Francis $f(x) = b'x + x'Ax$. This function is concave iff $-A$ is PSD, it is pseudo-concave (and thus quasi-concave) when $x'Ax$ is bilinear, i.e. A is of the form

$$\begin{bmatrix} 0 & \tilde{A} \\ \tilde{A}' & 0 \end{bmatrix}$$

We can write

$$f(x) = x'(b - Ax) = \sum x_i h_i(x),$$

with $h_i(x) = b_i - \sum a_{ij} x_j$. Consequently

$$u_i = \min_{x \in B} (b_i - \sum a_{ij} x_j).$$

Cabot & Francis use Murty's extreme point ranking procedure, but they do not seem to see that Murty's procedure of solving the fixed charge problem is in fact basically identical to the one they use (only applied to another type of function).

In the fixed charge problem

$$f(x) = \sum c_i x_i + \sum d_i \varepsilon_i$$

with

$$\varepsilon_i = \begin{cases} 0 & \text{if } x_i = 0 \\ 1 & \text{if } x_i > 0 \end{cases}$$

This can obviously be written in the form

$$f(x) = \sum x_i h_i(x)$$

with

$$h_i(x) = \begin{cases} c_i & \text{if } x_i = 0 \\ c_i + d_i & \text{if } x_i > 0 \end{cases}$$

and $g_B(x)$ is very easy to find. Murty already observed that this algorithm could very well be applied to $f(x) = \sum c_i x_i + C(x)$, with $C(x)$ a general concave function, by using $g_B(x) = \sum c_i x_i + G_0$, where G_0 is a lower bound of G on B. What we have done is merely relaxing concavity to quasi-concavity and combining this with the familiar cutting methods.

It obviously follows from this discussion that we must choose $g_B(x)$ from the convex set G which contains all linear functionals dominated by $f(x)$ on B in such a way that the lower bounds for $f(x)$ provided by $g_B(x)$ are as sharp as possible. We want (among other things)

$$\min_{x \in B} f(x) - \min_{x \in B} g_B(x)$$

to be as small as possible. This seems a rather difficult thing to demand, because we do not know f_{min}. If we have a set of linear functionals G, and

$$(\forall_G\, g_B)(\forall_B\, x): \; g_B(x) \leq f(x),$$

then we must solve

$$\max_{g_B \in G} \; \min_{x \in B} \; g_B(x).$$

If G is a finite set from which we can choose, this means that we must solve a finite sequence of LP-problems. In the second place we also want $g_B(x)$ to be such that

$$\max_{x \in B} g_B(x) \geq \min_{x \in B} f(x),$$

the more difference there is, the better. The reason for this requirement is clear. If $\max_{x \in B} g(x) \leq \min_{x \in B} f(x)$ our algorithm has to compute all vertices, and we do not gain much, compared with our previous proposals. If we select $g(x) = G_0$ for example, for all $x \in B$, then $\max_{x \in B} g_B(x) = G_0 \leq \min_{x \in B} f(x)$. If $\max g_B(x) = \min f(x)$, then the only advantage of our algorithm is that we do not have to investigate if there are any more vertices as soon as we have computed the last one. We also want g_B such that

$$\max_{x \in B} g_B(x) = \max_{g \in G} \; \max_{x \in B} g(x)$$

If $f(x)$ has some special structure, such as $H(\sum x_i h_i(x))$, then the cost of constructing a better function than $g_0(x) = \sum u_i x_i$ is probably much to high. If no linear functional is available in the first place, we might as well compute a good one.

We now turn to our own original (linearization) problem. For problem P we have

$$\begin{aligned} &\min\ \lambda_{min}(C), \\ &\quad Tr(B_k C) \geq 0, \\ &\qquad C \quad \text{PSD, SDC.} \end{aligned}$$

In the transformed version

$$\begin{aligned} f(x) = \lambda_{min}&(\textstyle\sum x_p S_p + Q) \quad \min\,! \\ &c'x \leq n(n-1), \\ &\quad x \geq 0, \\ &\quad \textstyle\sum x_p S_p + Q \qquad \text{PSD.} \end{aligned}$$

A possible linear function is easily found to be

$$g(x) = \sum x_p \lambda_{min}(S_p) + \lambda_{min}(Q),$$

and the problems Q_k are of the form

$$\begin{aligned} &\lambda_{min}(\textstyle\sum x_p S_p + Q) \quad \min! \\ &\left.\begin{aligned} &c'x \leq n(n-1), \\ &V_k x \leq b_k, \\ &\quad x \geq 0, \end{aligned}\right\} \text{ region } A_k \end{aligned}$$

where the matrix V_k has k rows. There is still some freedom in choosing the linear function, depending on which of the T_p matrices we eliminate. Let us consider the process of transforming from $\sum x_p T_p$ to $\sum x_p S_p + Q$ in some detail. We start with $C = \sum x_p T_p$, and $Tr(C) = \sum x_p Tr(T_p) = n(n-1)$. Solve for x_q

$$x_q = \frac{n(n-1) - \sum\limits_{p \neq q} x_p Tr(T_p)}{Tr(T_q)},$$

and substitute

$$C = \sum_{p \neq q} \left[T_p - \frac{Tr(T_p)}{Tr(T_q)} T_q \right] x_p + \frac{n(n-1)}{Tr(T_q)} T_q.$$

The matrices in the square brackets are the S_p. The added constraint is

$$\sum_{p \neq q} x_p \frac{\mathrm{Tr}(T_p)}{\mathrm{Tr}(T_q)} \leqslant \frac{n(n-1)}{\mathrm{Tr}(T_q)}.$$

If $\mathrm{Tr}(T_q) \geqslant 0$ then $x = 0$ is feasible for D_0, which means

$$\min_{x \in A_0} g(x) \leqslant \lambda_{min}(C) = \frac{n(n-1)}{\mathrm{Tr}(T_q)} \lambda_{min}(T_q) \leqslant \max_{x \in A_0} f(x).$$

If T_q is PSD, then

$$\max_{x \in A_0} g(x) \geqslant 0 > \min_{x \in A_0} f(x),$$

which is nice. If $T_q = q[nI - E]$, then

$$C = \sum_{p \neq q} \left[T_p - \frac{\mathrm{Tr}(T_p)}{n(n-1)} (nI - E) \right] x_p + (nI - E).$$

In this case the minimum eigenvalue of S_p is $\lambda_{min}(T_p) - \mathrm{Tr}(T_p)/(n-1) \leqslant 0$.
Of course $\lambda_{min}(C) = 0$, and consequently $g(x) \leqslant 0$ for all $x \in A_k$ (for all k).
It seems convenient to eliminate $nI - E$ (if present).

In our standard example we have

$$C = \begin{matrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{matrix}\, x_1 + \begin{matrix} 8 & -4 & -4 \\ -4 & 2 & 2 \\ -4 & 2 & 2 \end{matrix}\, x_2 + \begin{matrix} 4 & -5 & 1 \\ -5 & 4 & 1 \\ 1 & 1 & -2 \end{matrix}\, x_3$$

If we eliminate x_1 we get

$$\begin{matrix} 4 & -2 & -2 \\ -2 & -2 & 4 \\ -2 & 4 & -2 \\ & S_2 & \end{matrix} \qquad \begin{matrix} 2 & -4 & 2 \\ -4 & 2 & 2 \\ 2 & 2 & -4 \\ & S_3 & \end{matrix} \qquad \begin{matrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \\ & Q & \end{matrix}$$

with $\lambda_{min}(S_1) = \lambda_{min}(S_2) = -6$, and $\lambda_{min}(Q) = 0$. The added constraint is $2x_2 + x_3 \leqslant 1$, the LP-problem is

$$g(x_2, x_3) = -6x_2 - 6x_3 \quad \min!$$
$$2x_2 + x_3 \leqslant 1,$$
$$x_2,\ x_3 \geqslant 0.$$

vertices	$g(x_2,x_3)$	$\lambda_{min}(x_2,x_3)$
$(0,1)$	-6	-3
$(\frac{1}{2},0)$	-3	0
$(0,0)$	0	0

Elimination of x_2 gives

$$\begin{matrix} -6 & 3 & 3 \\ 3 & 3 & -6 \\ 3 & -6 & 3 \end{matrix} \qquad \begin{matrix} 0 & -3 & 3 \\ -3 & 3 & 0 \\ 3 & 0 & -3 \end{matrix} \qquad \begin{matrix} 4 & -2 & -2 \\ -2 & 1 & 1 \\ -2 & 1 & 1 \end{matrix}$$

$$S_1 \qquad\qquad S_3 \qquad\qquad Q$$

$$g(x_1,x_3) = -9x_1 - 3\sqrt{3}\,x_3 \quad \min !$$

$$3x_1 + x_3 \leqslant 1,$$

$$x_1,\ x_3 \geqslant 0.$$

vertices	$g(x_1,x_3)$	$\lambda_{min}(x_1,x_3)$
$(0,1)$	$-3\sqrt{3}$	-3
$(\frac{1}{3},0)$	-3	0
$(0,0)$	0	0

Finally elimination of x_3 gives

$$\begin{matrix} -6 & 12 & -6 \\ 12 & -6 & -6 \\ -6 & -6 & 12 \end{matrix} \qquad \begin{matrix} 0 & 6 & -6 \\ 6 & -6 & 0 \\ -6 & 0 & 6 \end{matrix} \qquad \begin{matrix} 4 & -5 & 1 \\ -5 & 4 & 1 \\ 1 & 1 & -2 \end{matrix}$$

$$S_1 \qquad\qquad S_2 \qquad\qquad Q$$

$$g(x_1,x_2) = -18x_1 - 6\sqrt{3}\,x_2 - 3$$

$$3x_1 + 2x_2 \leqslant 1,$$

$$x_1,\ x_2 \geqslant 0.$$

vertices	$g(x_1,x_2)$	$\lambda_{min}(x_1,x_2)$
$(\frac{1}{3},0)$	-9	0
$(0,\frac{1}{2})$	$-3 - 3\sqrt{3}$	0
$(0,0)$	-3	-3

Obviously elimination of x_3 gives the worst result (we have to compute all vertices). In the other cases we are ready after the computation of two vertices. Because $-3\sqrt{3} > -6$ elimination of x_2 seems best.

Appendix D: Corrections, addenda, remarks

D1: Definition 0.1 on page 1 is not correct. See appendices A and B.

D2: The proof of the fact that T is non-polyhedral in lemma 2.3 (page 9) is not too convincing. A more precise result is proved below.

Lemma 2.3: The set T of all symmetric PSD matrices is a nonpolyhedral pointed closed convex cone. The edges (minimal faces) of this cone are the symmetric PSD matrices of rank one.

Proof: That T is a convex cone is trivial. All PSD matrices of order n can be written as $\sum_{i=1}^{n} \lambda_i x_i x_i'$ (spectral decomposition) with $\lambda_i \geq 0$. Moreover a matrix of rank one cannot be written as a nonnegative linear combination of other PSD matrices, unless they are both proportional to the rank-one matrix itself. Consequently the matrices of type xx' are the edges. The cardinality of this set of edges equals the cardinality of the continuum, which means that T is not polyhedral. The fact that T is closed follows from the fact that it contains its minimal faces. Finally x'Cx = 0 for all real vectors x iff C = 0, which implies that T is pointed. Q.E.D.

D3: The phrase 'nonvoid interior' in corollary 2.3 (page 12) refers, of course, to the largest dimensional linear subspace in which the cone lies (and this subspace may even be two-dimensional, cf examples on p 40-42, p 45-47).

D4: Theorem 3.4 (page 22) is not quite correct. Inclusion is not necessarily strict, convergence in a finite number of steps can take place. Cf examples on page 40-42, 45-47. The critical point is that, although a finite number of the conditions $\mathrm{Tr}(y_s y_s' C) \geq 0$ is never sufficient for PSD, it may very well be true that a finite number of these conditions, taken together with $\mathrm{Tr}(E_{ij} C) \geq 0$, are indeed sufficient. This is exactly what happens in the examples mentioned previously. In corollary 3.3 we have $\lambda_s \leq 0$ for all s.

D5: It is not true that nI - W is always an edge. A counterexample can be found on page 47.

D6: It is useful to distinguish the <u>weak signature</u> λ^{w}_{ijkl} (defined on page 3) from the <u>semi-strong signature</u> defined by

$$\lambda^{s}_{ijkl} = \begin{cases} +1 & \text{if } \delta_{ij} >_{o} \delta_{kl}, \\ -1 & \text{if } \delta_{ij} <_{o} \delta_{kl}, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that $\delta_{ij} =_{o} \delta_{kl}$ implies $\lambda^{s}_{ijkl} = \lambda^{s}_{klij} = 1$, and that the system (1) on page 3 with λ^{s}_{ijkl} is equivalent to semi-strong monotonicity. In the formula for F on page 3 we may use either λ^{w}_{ijkl} or λ^{s}_{ijkl}, this does not make any difference.

D7: We are maximizing (on page 3-4)

$$F = \frac{\mathrm{Tr}(X'AX)}{\mathrm{Tr}(X'X)}.$$

If we take $X = K$ (the p eigenvectors corresponding with the p largest eigenvalues Λ), then

$$F_1 = \frac{\mathrm{Tr}(K'AK)}{\mathrm{Tr}(K'K)} = \frac{\sum \lambda_i}{p}.$$

But if we take $X = K\Lambda^{\frac{1}{2}}$, then

$$F_2 = \frac{\mathrm{Tr}(\Lambda^{\frac{1}{2}}K'AK\Lambda^{\frac{1}{2}})}{\mathrm{Tr}(\Lambda^{\frac{1}{2}}K'K\Lambda^{\frac{1}{2}})} = \frac{\sum \lambda_i^2}{\sum \lambda_i},$$

and by the CS-inequality $F_2 > F_1$. For F_2 the p largest eigenvalues must, of course, be nonnegative.

D8: At an earlier stage I planned appendices describing applications of basically the same method to other measurement models. But the paper is long enough already. I just mention them briefly in this short note. The method can be applied to the inner product model without much modifications. In fact all modifications that are needed are actually simplifications, although there are some new problems (B is a semi-strong solution, and in most cases $nI - B$ is, which implies $r_p = 0$). Representations theorems follow in about the same way. More interesting and more far-reaching modifications are needed to treat the 'city block' model. We work in the space of all real $n \times p$ matrices and we use a multidimensional variant of the Coombs trick. By considering all posible rank orderings of coordinate values we arrive at $(n!)^p$ polyhedral convex cones. The solution set is the union of these cones.

In the same way we can study the sup-metric, which is just a little bit more complicated. In the euclidean case we have to transform the whole problem into a different space, and we have to consider the solutions for all p n-1 together in order to get a convex (nonpolyhedral) problem. In the city block case we can study the problem for each p separately, we only work with (a large number of) polyhedral convex cones (linear inequalities). Of course the union of these cones is not convex. Moreover it may be empty for small p. In the linear nonmetric models (ACM, MFLP) we have a single convex polyhedral cone as the solution set (if a solution exists), but in most cases perfect solutions do not exist. We fist have to purify our system of inequalities, for example by finding the maximum solvable subset.

D9: p 64, top. Forget about appendix M.

D10: p 76. If T denotes the cone of PSD matrices then for each square symmetric matrix C_1

$$r_{min}(C_1) = \min_{C \in T} r(C, C_1) = \frac{\lambda_{min}(C_1)}{\sqrt{\sum \lambda_i^2(C_1)}} .$$

Each singular matrix has $r_{min} = 0$. Cf also theorem 2.5 on page 13.

D11: p 3-7. It is not sufficiently emphasized in this section that one of the main advantages of MRSCMS is that it generalizes quite easily to individual differences models in MDS and to special problems, like conditional matrices and unfolding situations.

D12: p 11, theorem 2.4, cf appendix B.

D13: p 17, lemma 3.1. The somewhat mysterious 'C T' in the first line of the proof means 'C must be a T-matrix'.

D14: p 21-25. It has not been sufficiently stressed in this section that we approximate the cone P both from the inside and from the outside. We have the sequence

$$Q_1 \supset Q_2 \supset \dots$$

with $Q_\infty = P$. On the other hand we may let S_k be the cone spanned by the PSD edges of Q_k, and then

$$S_1 \subset S_2 \subset \ldots$$

and again $S_\infty = P$. In the same way the value of r_p is also approximated from above and from below (which is quite important for evaluating convergence).

D15: p 13. As soon as we are working with closed and bounded convex sets an obvious alternative measure of uniqueness is the volume of the polyhedron. We get a sequence

$$v(\overline{Q}_0) \geqslant v(\overline{Q}_1) \geqslant v(\overline{Q}_2) \quad \ldots$$

coverging to $v(\overline{P})$. This volume has the attractive property that it is additve (in the sense that if we cut off T_i from $\overline{Q}_i$, then $v(\overline{Q}_{i+1}) = v(\overline{Q}_i) - v(T_i)$). Another attractive property is that we may study the volume of $\overline{Q}_i$ relative to a set $R \supset \overline{Q}_0$, and define

$$\overline{\pi}_i = \frac{v(\overline{Q}_i)}{v(R)},$$

which can obviously be interpreted as a probability measure. A possible choice is: $R = \overline{Q}_0$, which answers the question: if we sample C at random (uniform distribution) from the set of all scaled monotone matrices, what is the probability that is a solution, i.e. that it is PSD. We can also use for R the set of all scaled PSD matrices, and reverse the question. The definition corresponding with this uniqueness measure of a representative solution would be the centroid of $\overline{P}$.

D16: Appendix B. We have only given general representation theorems in this appendix, not representation theorems for special cases. The imbedding problem for one-dimensional euclidean scaling has a relatively simple solution, ans so has the imbedding problem for multidimensional unfolding. The solution of the space problem for E^1 simplifies for example to: a metric space is congruent with E^1 iff it is complete, convex, externally convex and it contains no equilateral triples. A metric space with more then four points in which for every triple of points one is metrically between the other two is congruently imbeddable in E_1.

D17: Appendix C, p 103. The expression for S_p at the bottom of page 103 shows that $Tr(S_p) = 0$. Because T_p and T_c cannot be proportional $S_p \neq 0$. Consequently $\lambda_{min}(S_p) < 0$, and $\lambda_{min}(S) \leqslant 0$ (because T_c is SDC). This means that $g(x) \leqslant 0$ for all $x \in A_k$, for all k, no matter what matrix T_σ we eliminate. Moreover $g(x) = 0$ iff $x = 0$ and T_σ is PSD.

References

Abelson, R.P. & Tukey, J.W.: Efficient conversion of non-metric information into metric information.
Proc. Amer. Statist. Assoc. Meetings, Social Sciences Section, 1959, 226-230.

Abelson, R.P & Tukey, J.W.: Efficient utilization of nonnumerical information in quantatitive analysis: general theory and the case of simple order.
Ann. Math. Statist., 1963, 34, 1347-1369.

Arrow, K.J., Hurwicz, L., & Uzawa, H.:
Studies in linear and nonlinear programming.
Stanford, Stanford Univ. Press, 1958.

Balinski, M.L.: An algorithm for finding all vertices of convex polyhedral sets.
J. Soc. Indust. and Appl. Math., 1961, 9, 72-88.

Beals, R. & Krantz, D.H.: Metrics and geodesics induced by order relations.
Math. Zeitschrift, 1967, 101, 285-298.

Beals, R., Krantz, D.H., & Tversky, A.:
Foundations of multidimensional scaling.
Psychol. Rev., 1968, 75, 127-143.

Berge, C.: Espaces topologiques, functions multivoques.
Paris, Dunod, 1966.

Blumenthal, L.M.: Distance geometries. A study of the development of abstract metrics.
Univ. of Missouri studies, 13, no 2, 1938.

Blumenthal, L.M.: Theory and applications of distance geometry.
Oxford, Clarendon Press, 1953.

Cabot, A.V. & Francis, R.L.: Solving certain nonconvex quadratic minimization

problems by ranking the extreme points.
Operations Research, 18, 1970, 82-96.

Cheney, E.W. & Goldstein, A.A.: Newton's method for convex programming and Tchebycheff approximation.
Numerische Math., 1, 1959, 253-265.

Collatz, L. & Wetterling, W.: Optimierungsaufgaben.
Berlin-Heidelberg-New York, Springer, 1966.

Cottle, R.W.: Symmetric dual quadratic programs.
Quart. Appl. Math., 21, 1963, 237-243.

Coombs, C.H.: A theory of data.
New York, Wiley, 1964.

Coombs, C.H.: Scaling and data theory.
In: Proc. NUFFIC international summer session on 'Psychological Measurement Theory', Den Haag, 1966, p 53-107.

Dantzig, G.B.: Linear programming and extensions.
Princeton, Princeton Univ. Press, 1963.

Dantzig, G.B. & Cottle, R.W.: Positive (semi-) definite programming.
In: Abadie, J. (ed): Nonlinear programming.
p 55-73. Amsterdam, North Holland, 1967.

De Leeuw, J.: Nonmetric multidimensional scaling.
Univ. of Leyden, Dep. of Data Theory,
Report RN 009-68, 1968.

De Leeuw, J.: The positive orthant method for nonmetric multidimensional scaling.
Univ. of Leyden, Dep. of Data Theory,
Report RN 001-70, 1970 (a).

De Leeuw, J.: The abstract structure of scaling theories. Univ. of Leyden, Dep. of Data Theory, Unpublished manuscript (as yet), 1970(b).

Dorn, W.S.: Self-dual quadratic programs. J. Soc. Ind. and Appl. Math., 1961, 9, 51-54.

Goldman, A.J.: Resolution and separation theorems for polyhedral convex sets. In: Kuhn, H.W. & Tucker, A.W. (eds), p 41-53

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

Hartley, H.O., & Hocking, R.R.: Convex programming by tangential approximation. Management Science, 9, 1963, 600-612.

Hartley, H.O., Hocking, R.R., & Cooke, W.P.: Least squares fit of definite quadratic forms by convex programming. Management Science, 13, 1967, 913-925.

Kelley, J.L. Jr.: The cutting plane method for solving convex programs. J. Soc. Indust. Appl. Math., 8, 1960, 703-712.

Krantz, D.H.: A survey of measurement theory. In: Dantzig, G.B. & Veinott, A.F. Jr. (eds): Mathematics of the decision sciences, part II, p 314-350, Providence, Amer. Math. Soc., 1968.

Kruskal, J.B.: Multidimensional scaling by optimizing goodness of-fit to a nonmetric hypotheses. Psychometrika, 1964, 29, 1-27.

Kruskal, J.B.: Nonmetric multidimensional scaling: a numeric method.
Psychometrika, 1964, 29, 115-129.

Kruskal, J.B., & Carroll, J. Douglas:
Geometrical models and badness-of-fit function
In: Krishnaiah, P.R. (ed): International symposium on multivariate analysis, Dayton, O
1966. New York, Academic Press, 1969.

Kuhn, H.W. & Tucker, A.W.: Nonlinear programming.
In: Neyman, J.: Proc. Sec. Berkeley Symp. on Math. Statist. and Prob., 481-492.
Berkeley, Univ. of California press, 1951.

Kuhn, H.W. & Tucker, A.W. (eds):
Linear inequalities and related systems.
Annals of mathematics studies 38, Princeton.
Princeton Univ. Press, 1956.

Ky Fan: Systems of linear inequalities.
In: Kuhn & Tucker (eds), p 99-156.

Lemke, C. E.: A method for solution of quadratic programs.
Management Science, 8, 1962, 442-453.

Lindman, H.R.: A numerical approximation to ordered metric scales.
Indiana Math. Psychol. Program, Report 69-2.
Dep. of Psychol., Indiana Univ., Bloomington, Indiana.

Lingoes, J.C.: Some boundary conditions for a monotone analys of symmetric matrices.
Michigan Math. Psychol. Program, Report MMPP 6
Dep. of Psychol., Univ. of Michigan, Ann Arbor Michigan.

Mangasarian, O.L.: Equilibrium points of bimatrix games. J. Soc. Indust. and Appl. Math., 1964, 12, 77

Mangasarian, O.L.: Nonlinear Programming. New York, McGraw-Hill, 1969.

Mangasarian, O.L. & Stone, H.: Two-person zero-sum games and quadratic programming. J. Math. Analysis and Appl., 1964, 9, 345-355.

Martos, B.: The direct power of adjacent vertex programming methods. Management Science, 1965, 12, 241-252.

Menger, K.: Untersuchungen über allgemeine Metrik. Math. Ann., 1928, 100, 75-163.

Motzkin, T.S., Raiffa, H., Thompson, G.L., & Thrall, R.M.: The double description method. In Kuhn, H.W. & Tucker, A.W. (eds): Contributions to the theory of games, II. Annals of Math. studies 28, p 51-73. Princeton, Princeton Univ. Press, 1953.

Murty, K.G.: Solving the fixed charge problem by ranking the extreme points. Operations Research, 1968, 16, 268-279.

Roskam, E.E.C.I.: A comparison of principles for algorithm construction in nonmetric scaling. Michigan Math. Psychol. Program, Report MMPP 6 Dep. of Psychol., Univ. of Michigan, Ann Arbor Michigan.

Roskam, E.E.C.I. Data theory and algorithms for nonmetric scali Parts I and II, mimeographed papers. Dep. of Math. Psychol., Catholic Univ. of

Nijmegen, Nijmegen, The Netherlands.

Schoenberg, I.J.: Remarks to Maurice Frechet's article "Sur la définition axiomatique d'une classe d'espaces vectorielles distanciés applicable vectoriellement sur l'espace de Hilbert".
Ann. of Math., 36, 1935, 724-732.

Shepard, R.N. & Caroll, J. Douglas:
Parametric representation of nonlinear data structures.
In: Krishnaiah, P.R. (ed): Int. Symp. on multivariate analysis, Dayton, Ohio, 1966.
New York, Academic Press, 1966.

Spence, I.: Nonmetric multidimensional scaling via direct search maximization of a nonanalytic criterion.
Mimeographed paper, Univ. of Toronto, 1969.

Sydow, H.: Ähnlichkeitsskalierung mittels allgemeiner Minkowskiräume.
Paper read at the 2nd annual meeting of the Gruppe Mathematische Psychologie in Marburg/La.
April 1968.

Taussky, O.: A recurring theorem on determinants.
Amer. Math. Monthly, 56, 1959, 1043-1044.

Tversky, A: Finite additive structures.
Michigan Math. Psychol. Program, Report MMP 6
Dep. of Psychol., Univ of Michigan, Ann Arbor, Michigan.

Tversky, A. & Krantz, D.H.: The dimensional representation and the metric structure of similarity data.
Michigan Math. Psychol. Program, Report 69-7

Dep. of Psychol, Univ. of Michigan, Ann Arbor, Michigan.

Uzawa, H.: A theorem on convex polyhedral cones. In Arrow et al (1958), p 23-31.

Uzawa, H.: An elementary method for linear programming. In Arrow et al (1958), p 179-188.

Veinott, A.F. Jr.: The supporting hyperplane method for unimodal programming. Operations Research, 1967, 15, 147-152.

Young, F.W.: A FORTRAN-IV program of nonmetric multidimensio scaling. Tech. Rep. 56, The L.L. Thurstone Psychometric Laboratory, Univ. of North Carolina, 1968.

Young, G. & Householder, A.S.: Discussion of a set of points in terms of thei mutual distances. Psychometrika, 1938, 3, 331-333.

Zangwill, W.I.: Nonlinear programming: a unified approach. Englewood Cliffs, Prentice-Hall, 1969.

Zoutendijk, G.: Methods of feasible directions: a study in linear and nonlinear programming. Amsterdam-London-New York-Princeton, Elsevier Publishing Co., 1960.