On Certain Linear Mappings Between Inner-Product and Squared-Distance Matrices

Frank Critchley
Department of Statistics
University of Warwick
Coventry CV4 7AL, England

Submitted by C. R. Rao

ABSTRACT

We obtain the spectral decomposition of four linear mappings. The first, κ, is a mapping of the linear hull of all centered inner-product matrices onto the linear hull of all the induced squared-distance matrices. It is based on the natural generalization of the cosine law of elementary Euclidean geometry. The other three mappings studied are κ^{-1}, the adjoint κ^*, and $(\kappa^*)^{-1}$. Extensions and applications, particularly to multidimensional scaling, are discussed in some detail.

1. INTRODUCTION

Consider a collection $X = \{x_i : i = 1, \dots, n\}$ of $n > 2$ points in a real inner-product space I. Suppose that X is centered (that is, $\Sigma x_i = 0$), and let $B_X \equiv (b_{ij})$ and $D_X \equiv (d_{ij})$ be the matrices defined by

$$b_{ij} = \langle x_i, x_j \rangle_I \quad \text{and} \quad d_{ij} = \|x_i - x_j\|_I^2,$$

where $\|\cdot\|_I$ is the norm on I induced by the inner product $\langle \cdot, \cdot \rangle_I$. Then it is clear that these two fundamental matrices are related by

$$d_{ij} = b_{ii} + b_{jj} - 2b_{ij} \tag{1κ}$$

and also, after a little algebra, by

$$b_{ij} = -\tfrac{1}{2}\left\{ d_{ij} - n^{-1}d_{i\cdot} - n^{-1}d_{\cdot j} + n^{-2}d_{\cdot\cdot} \right\} \tag{1τ}$$

with a dot denoting addition over an omitted subscript.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/88/$3.50

Let **B** and **D** denote respectively the sets of all such matrices B_X and D_X, and let **S** denote the vector space of all $n \times n$ real, symmetric matrices. Then (1) establishes a natural one-to-one correspondence between **B** and **D**, which we extend to the smallest subspaces of **S** containing them. Denoting these subspaces by $\mathbf{S}_C$ and $\mathbf{S}_H$ respectively, this extension of (1) is given by the pair of mappings $\kappa: \mathbf{S}_C \to \mathbf{S}_H$ and $\tau: \mathbf{S}_H \to \mathbf{S}_C$ defined by

$$\begin{aligned} \kappa(C) &= (C * I)1_n 1_n^T + 1_n 1_n^T (C * I) - 2C, \\ \tau(H) &= -\tfrac{1}{2}(I - J)H(I - J), \end{aligned} \tag{2}$$

where 1_n denotes the $n \times 1$ vector of ones, $J = n^{-1}1_n 1_n^T$, and $*$ is the Hadamard matrix product defined by $(A * Z)_{ij} = (A)_{ij}(Z)_{ij}$. We use the symbol κ to connote the fact that the first equation in (1) is just an application of the *cosine* law to the triangle with vertices 0, x_i, and x_j. The symbol τ is used in honour of Torgerson (1958), who gives an historical account of its derivation in the special case $I = \mathbb{R}^{n-1}$, $\langle x_i, x_j \rangle_I = x_i^T x_j$. The letter B is traditional in this case, while D connotes (squared) distance. We use C as a mnemonic for centered (zero row and column sums), and H as a mnemonic for hollow (zero diagonal entries).

We introduce the inner product on **S** defined by $\langle S_1, S_2 \rangle = \mathrm{tr}(S_1 S_2)$, which induces the Euclidean norm and its associated metric. Thus, **S** is a Hilbert space of dimension $m + n$, where $m \equiv \frac{1}{2}n(n-1)$, and is isomorphic to $\mathbf{E}^{(m+n)}$ in the obvious way. Similarly, each of its subspaces is a Hilbert space in the inherited inner product and is isomorphic to Euclidean space of the same finite dimensionality.

Observing that κ and τ are linear operators between Hilbert spaces, we introduce their adjoint operators $\kappa^*: \mathbf{S}_H \to \mathbf{S}_C$ and $\tau^*: \mathbf{S}_C \to \mathbf{S}_H$ defined by the relationships

$$\langle H, \kappa(C) \rangle = \langle \kappa^*(H), C \rangle \tag{3a}$$

and

$$\langle C, \tau(H) \rangle = \langle \tau^*(C), H \rangle, \tag{3b}$$

which are to hold for all H in $\mathbf{S}_H$ and for all C in $\mathbf{S}_C$.

The objective of this paper is to give an essentially complete account of these two pairs of operators by obtaining their spectral decompositions. This is done for κ and τ in Section 2. Their adjoints, for which we obtain explicit expressions, are dealt with in Section 3. Matrix representation of all four mappings are derived and studied in Section 4. Extensions and applications are discussed in the final section.

2. THE OPERATORS τ AND κ

The subspace $\mathbf{S}_C$ comprises all *centered* matrices, and $\mathbf{S}_H$ all *hollow* ones. That is:

PROPOSITION 2.1.

$$\mathbf{S}_C = \{C \in \mathbf{S} | C1_n = 0\} \quad \text{and} \quad \mathbf{S}_H = \{H \in \mathbf{S} | H * I = 0\}.$$

In particular, $\dim(\mathbf{S}_C) = m = \dim(\mathbf{S}_H)$.

Proof. For each $i < j$ let X_{ij} denote a collection in which $x_i = -x_j \neq 0$ while $x_k = 0$ for all $k \neq i$ or j. It is easy to see that the m members $D_{X_{ij}}$ of $\mathbf{D}$ are linearly independent and that the subspace $\{H \in \mathbf{S} | H * I = 0\}$ which they generate contains $\mathbf{D}$ and is the smallest subspace of $\mathbf{S}$ with this property. The proof for $\mathbf{S}_C$ is similar. ■

Throughout the paper, we use C and H to denote general members of $\mathbf{S}_C$ and of $\mathbf{S}_H$ respectively. The next result is of central importance.

THEOREM 2.2. *The mappings κ and τ are linear and mutually inverse.*

Proof. Linearity is immediate from (2). Suppose $H = \kappa(C)$. Noting that $(I - J)1_n = 0$ and $(I - J)C = C$, we have from the first equation in (2) that $(I - J)H(I - J) = -2C$. That is, $C = \tau(H)$. Conversely, suppose $C = \tau(H)$. Using $h_{ii} = 0$ and $c_{ij} = -\frac{1}{2}\{h_{ij} - n^{-1}h_{i.} - n^{-1}h_{.j} + n^{-2}h_{..}\}$ and simplifying, we find $c_{ii} + c_{jj} - 2c_{ij} = h_{ij}$. ■

The fact that κ and τ are mutually inverse is particularly convenient and is exploited as follows. Given a result about either $\mathbf{S}_H$ or $\mathbf{S}_C$, we at once write down an equivalent result for the other space. A result about H in $\mathbf{S}_H$ [C in $\mathbf{S}_C$] is immediately translated into one for C in $\mathbf{S}_C$ [H in $\mathbf{S}_H$] by writing $H = \kappa(C)$ [$C = \tau(H)$] and using $C = \tau(H)$ [$H = \kappa(C)$]. Pairs of such equivalent results are stated in {(a),(b)} form. Naturally, only one of them need be proved.

The spectral decompositions of the linear operators κ and τ are of fundamental interest. We say that two members of $\mathbf{S}$ are equivalent, written $\mathbf{S}_1 \sim \mathbf{S}_2$, if they have the same off-diagonal elements. Then, in a slight generalization of the usual definitions, we say that λ is an eigenvalue and H is an eigenmatrix of τ if $\tau(H) \sim \lambda H$ and $H \neq 0$. The members of the spectral decomposition of κ are similarly defined.

These spectral decompositions are obtained as follows. We define three subspaces of $\mathbf{S}$:

$$\mathbf{S}_{HC} = \mathbf{S}_H \cap \mathbf{S}_C, \qquad \mathbf{S}_J = \{aJ | a \in \mathbb{R}\}, \quad \text{and} \quad \mathbf{S}_W = \{w1_n^T + 1_n w^T | w^T 1_n = 0\}.$$

It is convenient to have a notation for the unique hollow and the unique centered matrix equivalent to a given symmetric matrix. Accordingly, we define $h: \mathbf{S} \to \mathbf{S}_H$ and $c: \mathbf{S} \to \mathbf{S}_C$ by

$$h(S) = S - (S * I) \quad \text{and} \quad c(S) = S - \{(S1_n 1_n^T) * I\}.$$

Since $\tau(H)$ is always centered, $\tau(H) \sim \lambda H$ if and only if $\tau(H) = \lambda c(H)$; since $\kappa(C)$ is always hollow, $\kappa(C) \sim \lambda C$ is equivalent to $\kappa(C) = \lambda h(C)$. Now the following three results can be verified by straightforward if tedious calculation. Insightful derivations, based on matrix representation of κ and τ, are given in Section 4.

PROPOSITION 2.3.

(a) *The subspaces* $h(\mathbf{S}_J)$, $h(\mathbf{S}_W)$, *and* $\mathbf{S}_{HC}$ *are pairwise orthogonal and have direct sum* $\mathbf{S}_H$. *Their dimensionalities are* 1, $n-1$, *and* $m-n$ *respectively.*

(b) *The subspaces* $c(\mathbf{S}_J)$, $c(\mathbf{S}_W)$, *and* $\mathbf{S}_{HC}$ *are pairwise orthogonal and have direct sum* $\mathbf{S}_C$. *They have dimensions* 1, $n-1$, *and* $m-n$ *respectively.*

Observe that $m-n$ vanishes if and only if $n=3$. In this case $\mathbf{S}_{HC}$ is the trivial subspace.

THEOREM 2.4.

(a) (i) *On* $h(\mathbf{S}_J)$, $\tau(H) \sim (-2n)^{-1}H$;
(ii) *on* $h(\mathbf{S}_W)$, $\tau(H) \sim (-n)^{-1}H$;
(iii) *on* $\mathbf{S}_{HC}$, $\tau(H) = -\frac{1}{2}H$.

(b) (i) *On* $c(\mathbf{S}_J)$, $\kappa(C) \sim (-2n)C$;
(ii) *on* $c(\mathbf{S}_W)$, $\kappa(C) \sim (-n)C$;
(iii) *on* $\mathbf{S}_{HC}$, $\kappa(C) = (-2)C$.

Let the mappings α_J and β_J represent orthogonal projection of $\mathbf{S}_H$ onto $h(\mathbf{S}_J)$ and of $\mathbf{S}_C$ onto $c(\mathbf{S}_J)$ respectively. Let α_W, α_{HC} and β_W, β_{HC} be similarly defined. We denote $\alpha_J(H)$ by H_J and $\beta_J(C)$ by C_J, with similar conventions for H_W, H_{HC} and C_W, C_{HC}.

PROPOSITION 2.5. *We have the orthogonal decompositions*

(a) $H = H_J + H_W + H_{HC}$ and
(b) $C = C_J + C_W + C_{HC}$,

where

(a) (i) $H_J = n(n-1)^{-1}h(JHJ)$,
(ii) $H_W = n(n-2)^{-1}h\{(I-J)HJ + HJ(I-J)\}$,
(iii) $H_{HC} = H - n(n-2)^{-1}h(HJ + JH) - n^2(n-1)^{-1}(n-2)^{-1}h(JHJ)$,

and

(b) (i) $C_J = (n-1)^{-1}\mathrm{tr}(C)(I-J)$,
(ii) $C_W = -(n-2)^{-1}\{c(c_*1_n^T + 1_n c_*^T) + 2\,\mathrm{tr}(C)(I-J)\}$,
(iii) $C_{HC} = C + (n-2)^{-1}c(c_*1_n^T + 1_n c_*^T) + n(n-1)^{-1}(n-2)^{-1}\mathrm{tr}(C)(I - J)$,

with

$$c_* \equiv (C * I)1_n.$$

As the mappings κ and τ are linear, we obtain their spectral decompositions by combining Proposition 2.3 and Theorem 2.4, while Proposition 2.5 provides the explicit form of the six orthogonal projections $\alpha_J, \ldots, \beta_{HC}$. Let $\circ$ denote composition of maps. We have then the desired result:

THEOREM 2.6.

(i) (a) $\tau = (-2n)^{-1}(c \circ \alpha_J) + (-n)^{-1}(c \circ \alpha_W) + (-2)^{-1}\alpha_{HC}$,
(b) $\kappa = (-2n)(h \circ \beta_J) + (-n)(h \circ \beta_W) + (-2)\beta_{HC}$.

(ii) *That is, for all H in $\mathbf{S}_H$ and for all C in $\mathbf{S}_C$,*
(a) $\tau(H) = (-2n)^{-1}c(H_J) + (-n)^{-1}c(H_W) + (-2)^{-1}H_{HC}$,
(b) $\kappa(C) = (-2n)h(C_J) + (-n)h(C_W) + (-2)C_{HC}$.

COROLLARY 2.7. *The zero matrix is the unique fixed point of both τ and κ.*

Proof. Immediate. ∎

COROLLARY 2.8. *The relevant restrictions of* τ *and* κ *are mutually inverse linear mappings between*

(i) $h(\mathrm{S}_J)$ *and* $c(\mathrm{S}_J)$,
(ii) $h(\mathrm{S}_W)$ *and* $c(\mathrm{S}_W)$,
(iii) S_{HC} *and itself.*

Proof. Observe that $c \circ h = c$ and $h \circ c = h$. ■

COROLLARY 2.9.

(a) (i) *On* $h(\mathrm{S}_J)$, $\|\tau(H)\| = (2\sqrt{n})^{-1}\|H\|$,
(ii) *on* $h(\mathrm{S}_W)$, $\|\tau(H)\| = (\sqrt{2n})^{-1}\|H\|$,
(iii) *on* S_{HC}, $\|\tau(H)\| = \frac{1}{2}\|H\|$;

(b) (i) *on* $c(\mathrm{S}_J)$, $\|\kappa(C)\| = 2\sqrt{n}\,\|C\|$,
(ii) *on* $c(\mathrm{S}_W)$, $\|\kappa(C)\| = \sqrt{2n}\,\|C\|$,
(iii) *on* S_{HC}, $\|\kappa(C)\| = 2\|C\|$.

Proof. On $h(\mathrm{S}_J)$, $\tau = (-2n)^{-1}c$ and $\|c(H)\| = \sqrt{n}\,\|H\|$, yielding a(i). The rest of the proof is similar. ■

Finally, using the induced operator norm, let

$$\|\tau\| = \sup\left\{ \|\tau(H)\| \,\middle|\, \|H\| \leqslant 1 \right\}.$$

Then, omitting for brevity the special case $n = 3$ where $\mathrm{S}_{HC} = \{0\}$, we obtain:

COROLLARY 2.10. *Suppose* $n > 3$, *and let* $0 \neq H \in \mathrm{S}_H$ *and* $0 \neq C \in \mathrm{S}_C$.

(a) $(2\sqrt{n})^{-1}\|H\| \leqslant \|\tau(H)\| \leqslant \frac{1}{2}\|H\|$, *with equality in the first place if and only if* $H \in h(\mathrm{S}_J)$, *and in the second if and only if* $H \in \mathrm{S}_{HC}$.

(b) $2\|C\| \leqslant \|\kappa(C)\| \leqslant (2\sqrt{n})\|C\|$, *with equality in the first place if and only if* $C \in \mathrm{S}_{HC}$ *and in the second if and only if* $C \in c(\mathrm{S}_J)$.

In particular, $\|\tau\| = \frac{1}{2}$ *and* $\|\kappa\| = 2\sqrt{n}$.

Proof. Immediate. ■

3. THE OPERATORS τ^* AND κ^*

By Theorem 2.2 and Corollary 2.10, τ and κ are mutually inverse, bounded linear operators between the Hilbert spaces $\mathbf{S}_H$ and $\mathbf{S}_C$. By the general theory of adjoints [see, for example, Luenberger, (1969, §§ 6.5–6.8)], the operators $\kappa^*: \mathbf{S}_H \to \mathbf{S}_C$ and $\tau^*: \mathbf{S}_C \to \mathbf{S}_H$ defined by (3) have therefore the properties stated in the following portmanteau theorem. Recall that the polar $\mathbf{A}^\circ(\mathbf{L})$, or $\mathbf{A}^\circ$ for brevity, of a subset $\mathbf{A}$ of $\mathbf{S}$ relative to a subspace $\mathbf{L}$ of $\mathbf{S}$ is defined by

$$\mathbf{A}^\circ(L) = \{L \in \mathbf{L} | \text{for all } A \text{ in } \mathbf{A}, \langle L, A \rangle \leqslant 0\}.$$

We denote the inverse image of a set by the superscript (-1).

THEOREM 3.1.

(i) *κ^* and τ^* are bounded, linear, and mutually inverse.*

(ii) (a*) $\|\kappa^*\| = \|\kappa\|$.
(b*) $\|\tau^*\| = \|\tau\|$.

(iii) (a*) $\tau^{**} = \tau$.
(b*) $\kappa^{**} = \kappa$.

(iv) (a*) *For all* $\mathbf{A} \subset \mathbf{S}_C$, $\{\kappa(\mathbf{A})\}^\circ = \kappa^{*(-1)}(\mathbf{A}^\circ)$, *both polars being relative to* $\mathbf{S}_C$.
(b*) *For all* $\mathbf{A} \subset \mathbf{S}_H$, $\{\tau(\mathbf{A})\}^\circ = \tau^{*(-1)}(\mathbf{A}^\circ)$, *both polars being relative to* $\mathbf{S}_H$.

We proceed by direct analogy with our treatment of τ and κ in Section 2. In particular, equivalent results, reflecting now the one-to-one correspondence which κ^* and τ^* provide between $\mathbf{S}_H$ and $\mathbf{S}_C$, are given as $\{(a^*), (b^*)\}$ pairs. The following explicit expressions for κ^* and τ^* may be verified directly. More insightful derivations are given at the end of Section 4.

THEOREM 3.2.

(a*) $\kappa^* = -2c_H$,
(b*) $\tau^* = -\frac{1}{2}h_C$,

where c_H and h_C are the restrictions of c and h to $\mathbf{S}_H$ and $\mathbf{S}_C$ respectively.

That is,

(a*) $\kappa^*(H) = -2\{H - \mathrm{diag}(H1_n)\}$,
(b*) $\tau^*(C) = -\frac{1}{2}\{C - (C * I)\}$.

The spectral decompositions of κ^* and τ^* are simple and immediate.

THEOREM 3.3.

(a*) $\kappa^*(H) \sim (-2)H$ *on* $\mathbf{S}_H$.
(b*) $\tau^*(C) \sim (-\frac{1}{2})C$ *on* $\mathbf{S}_C$.

COROLLARY 3.4. *The zero matrix is the unique fixed point of both* κ^* *and* τ^*.

Proof. Immediate. ■

COROLLARY 3.5. *Let* $\mathbf{L}$ *denote* any *subspace of* $\mathbf{S}$. *Then the relevant restrictions of* κ^* *and* τ^* *are mutually inverse linear mappings between* $h(\mathbf{L})$ *and* $c(\mathbf{L})$.

Proof. Immediate. ■

COROLLARY 3.6.

(a*) (i) *On* $h(\mathbf{S}_I)$, $\|\kappa^*(H)\| = (2\sqrt{n})\|H\|$;
(ii) *on* $h(\mathbf{S}_W)$, $\|\kappa^*(H)\| = (\sqrt{2n})\|H\|$;
(iii) *on* $\mathbf{S}_{HC}$, $\|\kappa^*(H)\| = 2\|H\|$.

(b*) (i) *On* $c(\mathbf{S}_I)$, $\|\tau^*(C)\| = (2\sqrt{n})^{-1}\|C\|$;
(ii) *on* $c(\mathbf{S}_W)$, $\|\tau^*(C)\| = (\sqrt{2n})^{-1}\|C\|$;
(iii) *on* $\mathbf{S}_{HC}$, $\|\tau^*(C)\| = \frac{1}{2}\|C\|$.

Proof. Combine Theorem 3.3 with Corollary 2.9. ■

Omitting again the special case $n = 3$ for brevity, we obtain at once:

COROLLARY 3.7. *Suppose* $n > 3$, *and let* $0 \neq H \in \mathbf{S}_H$ *and* $0 \neq C \in \mathbf{S}_C$.

(a*) $2\|H\| \leq \|\kappa^*(H)\| \leq (2\sqrt{n})\|H\|$, *with equality in the first place if and only if* $H \in \mathbf{S}_{HC}$, *and in the second if and only if* $H \in h(\mathbf{S}_I)$.

(b*) $(2\sqrt{n})^{-1}\|C\| \leq \|\tau^*(C)\| \leq \frac{1}{2}\|C\|$, *with equality in the first place if and only if* $C \in c(\mathbf{S}_I)$, *and in the second if and only if* $C \in \mathbf{S}_{HC}$.

Let τ_J and $(\tau^*)_J$ denote the restrictions of τ and τ^* to $h(\mathbf{S}_J)$ and $c(\mathbf{S}_J)$ respectively. Let τ_W, $(\tau^*)_W$, τ_{HC}, $(\tau^*)_{HC}$, and their analogues for κ and κ^* be similarly defined. Then combining Corollaries 2.8 and 3.5, we obtain:

PROPOSITION 3.8.

(i) $(\tau_J)^* = (\tau^*)_J$ *and* $(\kappa_J)^* = (\kappa^*)_J$.
(ii) $(\tau_W)^* = (\tau^*)_W$ *and* $(\kappa_W)^* = (\kappa^*)_W$.
(iii) $(\tau_{HC})^* = (\tau^*)_{HC}$ *and* $(\kappa_{HC})^* = (\kappa^*)_{HC}$.

Verbally, this last result says that for both τ and κ, and for each of their three eigenspaces, the restriction of the adjoint is the adjoint of the restriction.

4. MATRIX REPRESENTATIONS OF τ AND κ

For S in $\mathbf{S}$, let $v(S) \in \mathbb{R}^m$ contain the above-diagonal elements of S listed in rowwise order. That is, $v(S) = (s_{12}, \ldots, s_{1n}, s_{23}, \ldots, s_{n-1,n})^T$. Then $H \leftrightarrow v(H)$ and $C \leftrightarrow v(C)$ provide natural one-to-one correspondences from $\mathbf{S}_H$ and $\mathbf{S}_C$ to $\mathbb{R}^m$. We use these correspondences to represent τ and κ by $m \times m$ matrices T and K via

(a) $\tau(H) = C \leftrightarrow Tv(H) = v(C)$ and
(b) $\kappa(C) = H \leftrightarrow Kv(C) = v(H)$.

In particular, $\tau(H) \sim \lambda H$, $H \neq 0$, if and only if $v(H)$ is an eigenvector of T corresponding to the eigenvalue λ. Our first task is then to obtain the spectral decompositions (in the usual sense) of T and K.

We obtain explicit expressions for T and K as follows. For $p \geqslant 1$, let $\mathbf{W}_p$ be the $(p-1)$-dimensional subspace of $\mathbb{R}^p$ defined by $\mathbf{W}_p = \{w \in \mathbb{R}^p | w^T 1_p = 0\}$. With A as a mnemonic for "additive," let $\mathbf{S}_A$ be the n-dimensional subspace of $\mathbf{S}$ defined by $\mathbf{S}_A = \{x 1_n^T + 1_n x^T | x \in \mathbb{R}^n\}$. Finally, define the $m \times n$ binary matrix R by

$$R^T = \begin{bmatrix} 1_{n-1}^T & 0^T & \cdots & 0^T & 0 \\ & 1_{n-2}^T & \cdots & 0^T & 0 \\ & & & \vdots & \vdots \\ & & & 0^T & 0 \\ I_{n-1} & & & & \\ & I_{n-2} & & 1_2^T & 0 \\ & & \cdots & \boxed{I_2} & \begin{matrix} 1 \\ 1 \end{matrix} \end{bmatrix},$$

where the suffix on I (as on J below) denotes its order. This matrix R is extremely useful. Its properties are summarized in the following result, whose straightforward proof we omit.

PROPOSITION 4.1.

(i) *R has full column* rank n.

(ii) *For all* $x \in \mathbb{R}^n$, $Rx = v(x1_n^T + 1_n x^T)$. *In particular,* $R1_n = 2 \cdot 1_m$, *while* $Rx \in \mathbf{W}_m$ *if and only if* $x \in \mathbf{W}_n$.

(iii) *For all* $S \in \mathbf{S}$, $R^T v(S)$ *has ith element* $(s_{i\cdot} - s_{ii})$. *In particular,* $R^T 1_m = (n-1)1_n$, *while* $R^T v(S) \in \mathbf{W}_n$ *if and only if* $v(S) \in \mathbf{W}_m$.

(iv) *The common range space of R and of RR^T is $v(\mathbf{S}_A)$. The common null space of R^T and of RR^T is $v(\mathbf{S}_{HC})$.*

(v) $R^T R = (n-2)I_n + nJ_n$.

Next, we observe that τ can be expressed in terms of the row sums of H, and κ in terms of the diagonal entries of C. Accordingly, we define h_+ and c_* in $\mathbb{R}^n$ by

$$h_+ = H1_n \quad \text{and} \quad c_* = (C * I)1_n$$

and rewrite (2) as:

PROPOSITION 4.2.

(a) $\tau(H) = -\frac{1}{2}\{H - n^{-1}h_+ 1_n^T - n^{-1}1_n h_+^T + n^{-1}(1_n^T h_+)J_n\}$,

(b) $\kappa(C) = c_* 1_n^T + 1_n c_*^T - 2C$.

Combining Propositions 4.1 and 4.2, we find, as required, that

THEOREM 4.3.

(a) $T = -\frac{1}{2}\{I_m - n^{-1}RR^T + (1 - n^{-1})J_m\}$,

(b) $K = -(2I_m + RR^T)$.

Proof. Observe that $h_+ = R^T v(H)$ and $c_* = -R^T v(C)$. ■

COROLLARY 4.4. *T and K are mutually inverse and symmetric, negative definite.*

Proof. T and K are mutually inverse because τ and κ are so (Theorem 2.2). It is clear from Theorem 4.3(b) that K is symmetric, negative definite, and so therefore is K^{-1}. ■

We obtain the spectral decompositions of T and K from that of RR^T as follows. Here orthogonality is with respect to the usual inner product $\langle y, z \rangle = y^T z$ on $\mathbb{R}^m$.

PROPOSITION 4.5. $\mathbb{R}^m = v(\mathbf{S}_J) \oplus v(\mathbf{S}_W) \oplus v(\mathbf{S}_{HC})$ *is an orthogonal decomposition in which the subspaces stated have dimensions* 1, $n-1$, *and* $m-n$ *respectively.*

Proof. Straightforward. ∎

Let P_J denote the matrix projecting $\mathbb{R}^m$ orthogonally onto $v(\mathbf{S}_J)$, and let P_W, P_{HC} be similarly defined. For y in $\mathbb{R}^m$, write $y_J = P_J y$, and let y_W, y_{HC} be similarly defined.

PROPOSITION 4.6. *Any vector y in $\mathbb{R}^m$ can be decomposed orthogonally as*

$$y = y_J + y_W + y_{HC},$$

where $y_J = m^{-1}(1_m^T y)1_m$ *and* $y_W = (n-2)^{-1} v(w 1_n^T + 1_n w^T)$, $w = (I_n - J_n)R^T y$.

Proof. Consider the two matrices J_m and $(n-2)^{-1}R(I_n - J_n)R^T$. The first is clearly symmetric and idempotent, and so, by Proposition 4.1, is the second. They therefore represent orthogonal projection onto their range spaces. But $\text{Range}(J_m) = v(\mathbf{S}_J)$. Thus $P_J = J_m$, and so y_J is as stated. Now $I_n - J_n$ is the symmetric, idempotent matrix projecting $\mathbb{R}^n$ onto $\mathbf{W}_n$. Thus, using the fact that $\text{Range}(A) = \text{Range}(AA^T)$, the second matrix has range space $\{Rw \mid w \in \mathbf{W}_n\}$. But, by Proposition 4.1(ii), this is $v(\mathbf{S}_W)$. Thus $P_W = (n-2)^{-1}R(I_n - J_n)R^T$ and so, using Proposition 4.1(ii) again, y_W is as stated. Finally, Proposition 4.5 implies that $P_J + P_W + P_{HC} = I_m$ and that $P_J P_W = P_W P_{HC} = P_{HC} P_J = 0$. ∎

PROPOSITION 4.7. RR^T *has spectral decomposition* $RR^T = (2n-2)P_J + (n-2)P_W$.

Proof. By Proposition 4.1(v), $R^T R$ has spectral decomposition

$$R^T R = (2n-2)J_n + (n-2)(I_n - J_n).$$

That is, $R^TR1_n = (2n-2)1_n$ and $R^TRw = (n-2)w$ for all $w \in \mathbf{W}_n$. Thus, using Proposition 4.1(ii), $RR^T1_m = (2n-2)1_m$ and $RR^Tw = (n-2)w$ for all $w \in \mathbf{W}_m$. ∎

THEOREM 4.8. *The spectral decompositions of T and K are as follows*:

(a) $T = (-2n)^{-1}P_J + (-n)^{-1}P_W + (-2)^{-1}P_{HC}$,
(b) $K = (-2n)P_J + (-n)P_W + (-2)P_{HC}$.

In particular,

(i) *on* $v(\mathbf{S}_J)$, $Ty = (-2n)^{-1}y$ *and* $Ky = (-2n)y$;
(ii) *on* $v(\mathbf{S}_W)$, $Ty = (-n)^{-1}y$ *and* $Ky = (-n)y$;
(iii) *on* $v(\mathbf{S}_{HC})$, $Ty = (-2)^{-1}y$ *and* $Ky = (-2)y$.

Proof. The spectral decomposition of K follows from combining Propositions 4.5 and 4.7 with Theorem 4.3(b). That of $T = K^{-1}$ is then immediate. ∎

We now translate these results about T and K back into properties of τ and κ. To effect this, let $\mathbf{H}$ and $\mathbf{C}$ denote $\mathbb{R}^m$ endowed respectively with the inner products

$$\langle y, z\rangle_{\mathbf{H}} = 2y^Tz \quad \text{and} \quad \langle y, z\rangle_{\mathbf{C}} = y^T(-K)z.$$

THEOREM 4.9. *The correspondences* $H \leftrightarrow v(H)$ *and* $C \leftrightarrow v(C)$ *establish Hilbert-space isomorphisms between* $\mathbf{S}_H$ *and* $\mathbf{H}$ *and between* $\mathbf{S}_C$ *and* $\mathbf{C}$ *respectively*.

Proof. The result for $\mathbf{S}_H$ is immediate. That for $\mathbf{S}_C$ follows on recalling the explicit form of K [Theorem 4.3(b)] and that $c_* = -R^Tv(C)$. ∎

THEOREM 4.10. *Let y and z be eigenvectors of K (or T) corresponding to distinct eigenvalues. Then* $\langle y, z\rangle_{\mathbf{H}} = 0 = \langle y, z\rangle_{\mathbf{C}}$.

Proof. Observe that $y^Tz = y^TKz = 0$. ∎

Combining these two theorems, Proposition 4.5 translates into Proposition 2.3, and Proposition 4.6 into Proposition 2.5. Theorem 2.4 is the translation of the latter part of Theorem 4.8.

Next we discuss matrix representations of κ^* and τ^*. As with κ and τ, we use the correspondences $H \leftrightarrow v(H)$ and $C \leftrightarrow v(C)$ to represent κ^* and

τ^* by $m \times m$ matrices K^* and T^* via:

(a*) $\kappa^*(H) = C \leftrightarrow K^*v(H) = v(C)$ and
(b*) $\tau^*(C) = H \leftrightarrow T^*v(C) = v(H)$.

THEOREM 4.11.

(a*) $K^* = -2I_m$.
(b*) $T^* = -\frac{1}{2}I_m$.

Proof. Using the defining relation (3a) and Theorem 4.9 gives

$$2\{v(H)\}^T Kv(C) = \{K^*v(H)\}^T(-K)v(C)$$

for all H in $\mathbf{S}_H$ and for all C in $\mathbf{S}_C$. Thus $K^* = -2I_m$. By Theorem 3.1(i), $\tau^* = (\kappa^*)^{-1}$ and so $T^* = (K^*)^{-1}$. ∎

Finally, we observe that the simplicity of the adjoints κ^* and τ^* results from the rather remarkable fact that the matrix defining the inner product on **C** is the negative of the matrix representing κ, while that defining the inner product on **H** is just a multiple of the identity.

5. DISCUSSION

5.1. *Extensions*

Let $\mathbf{B}^\circ$ and $\mathbf{D}^\circ$ denote the polars of **B** and **D** with respect to their linear hulls $\mathbf{S}_C$ and $\mathbf{S}_H$ respectively. By Theorem 2.2, $\mathbf{D} = \kappa(\mathbf{B})$ and $\mathbf{B} = \tau(\mathbf{D})$. By Theorem 3.1, $\mathbf{B}^\circ = \kappa^*(\mathbf{D}^\circ)$ and $\mathbf{D}^\circ = \tau^*(\mathbf{B}^\circ)$. Now **B** is self-dual. That is, $\mathbf{B}^\circ = -\mathbf{B}$. Thus, each of the four sets **B**, **D**, $\mathbf{B}^\circ$, and $\mathbf{D}^\circ$ is a known nonsingular linear transformation of each of the others. In a companion paper (Critchley, 1986a), we use the four mappings studied in the present paper to derive in a unified way the properties of these four sets and to illuminate our understanding of the relationships between these properties. For example, being a pointed solid closed convex cone is a property *invariant* to nonsingular linear transformation. Having established this property for **B**, it is then immediate that **D**, $\mathbf{B}^\circ$, and $\mathbf{D}^\circ$ all share it. Again, support cones are *equivariant* to nonsingular linear transformation. Thus, having obtained the support cone at a boundary point of **B**, we have at once by transformation the support cone at the corresponding boundary point of each of **D**, $\mathbf{B}^\circ$, and $\mathbf{D}^\circ$.

In a related paper (Critchley, 1986b), we study the behavior of the rank and spectral decomposition of a matrix under the mappings τ, κ and their adjoints.

There are close links with the recent work on diversity and quadratic entropy reported in Rao (1982, 1984, 1986). In particular, Equation (4.4) of Rao (1984) involves an extension of the domain of κ from $\mathbf{S}_C$ to $\mathbf{S}$.

We now briefly note some of the other possible extensions to the present paper. First, as is natural in some contexts, we may allow more general and possibly different inner products on $\mathbf{S}_H$ and $\mathbf{S}_C$. This leaves the mappings τ and κ unchanged, but not their adjoints. Secondly, we may extend the domain of both τ and κ to the space of all real $n \times n$ matrices, with consequent changes to their adjoints. This throws light on the analysis of nonsymmetric data. Extending further to the space and all real $l \times n$ matrices, we recover the removal of row and column effects in statistical models for two-way data as the transformation $-2\tau(\cdot)$. Thirdly, following Gower (1982), we may consider generalizations of the normalization condition $\Sigma x_i = 0$. This is a more fundamental change, which affects $\mathbf{B}$ and each of the four mappings studied, but not the set $\mathbf{D}$. Finally, we may envisage extending the collection $\{x_i : i = 1, \dots, n\}$ to a countable infinity or continuum of points. Clearly this requires restriction or normalization of the mappings τ and κ in some sense to prevent their blowing up as the number of points increases.

5.2. *Applications*

Our interest in the present paper arose out of multidimensional scaling. For an excellent introduction to this subject see Kruskal and Wish (1978), and for a recent authoritative review of its theory and algorithms see De Leeuw and Heiser (1982). In the multidimensional scaling context, we identify a collection X with an $n \times p$ *configuration matrix* $\tilde{X}$ of n points in $I = \mathbb{R}^p$ for some $p \leqslant (n-1)$. Depending upon the particular application on hand, it is often appropriate to measure the squared distance between the ith and jth points by their squared Euclidean distance, or by a weighted version $(g_i x_i - g_j x_j)^T (g_i x_i - g_j x_j)$ of this in which $g_i > 0$ reflects the importance of the ith point in some sense, or by $(x_i - x_j)^T \operatorname{diag}(\tilde{g})(x_i - x_j)$ in which the positive elements of the vector $\tilde{g}$ reflect the relative importances of the dimensions, or, most generally of all, by a squared Mahalanobis distance $(x_i - x_j)^T M (x_i - x_j)$ for some symmetric, positive definite M. Each of these possibilities, and several others, can be accommodated in the present paper by an appropriate choice of the inner product on I and/or an initial transformation (such as $x_i \to g_i x_i$).

Taken together, the present paper and its companion (Critchley, 1986a) provide a mathematical framework with many fruitful applications to the study of multidimensional scaling and related methods of data analysis. In particular, it provides a means of their clear and precise comparison. For example, Critchley (1980) uses this framework to characterize nine such methods as contrasting optimization problems. This is useful (1) in giving theoretical insights into these methods, (2) in establishing their formal properties, and (3) in devising algorithms for their implementation. In Critchley (1986c) we focus upon two particular methods and use them to illustrate each of these three aspects in turn. The theme of Critchley (1980) is that a variety of data analysis problems can be posed as projection onto a closed convex cone. This establishes at once the existence and uniqueness of a solution to the problem. Moreover the solution can be simply characterized in terms of three conditions. For example, squared-distance multidimensional scaling is projection of a given H onto $\mathbf{D}$. The unique solution $\hat{D}$ is characterized by $\hat{D} \in \mathbf{D}$, $\langle H - \hat{D}, \hat{D} \rangle = 0$, and thirdly $(H - \hat{D}) \in \mathbf{D}^{\circ} = \tau^{*}(-\mathbf{B})$. Hence the importance of studying adjoints and polars. In a review paper (Critchley, 1986d), we use this framework to unify and extend the literature on Euclidean dimensionality theorems in multidimensional scaling and hierarchical cluster analysis. In particular, certain results reported in Lingoes (1971), Holman (1972), De Leeuw and Heiser (1982), and the present author's D. Phil. thesis are generalized or unified there.

Gower (1982, 1984, 1985) has initiated a study of the theory of distance matrices. See also Mathar (1985). In particular, Gower shows that $HH^{-}1_n = 1_n$ for any generalized inverse of any nonzero H in $\mathbf{D}$. In a review paper (Gower, 1986), he raises the problem of finding additional requirements on a matrix H in $\mathbf{S}_H$ which are necessary and sufficient for $H \in \mathbf{D}$. A solution to this problem is given in Critchley (1986b) based on certain extensions to the present paper.

One statistical approach to multidimensional scaling is via a probability model for a dissimilarity matrix or, more generally, for H in $\mathbf{S}_H$. A collection $\{h_{ij} : 1 \leqslant i < j \leqslant n\}$ of random variables identifies a matrix H in the obvious way. Consider the probability model $H = D + U$ in which $D \in \mathbf{D}$ is an unknown true matrix about which we wish to make inferences based on a value of H observed in the presence of errors U. Let $u = v(U)$, $\mu = E(u)$, and $\Omega_H = \operatorname{cov}(u) = \operatorname{cov}(v(H))$, where $\operatorname{cov}(\cdot)$ denotes a covariance matrix. Often we will take $\mu = 0$. Then $C \equiv \tau(H)$ is the true inner-product matrix $B \equiv \tau(D)$ plus the matrix $V \equiv \tau(U)$ of errors. By linearity, $v(V) = Tu$ and thus is multivariate normal if and only if u is so. Moreover, in the general case, the properties of $E\{v(V)\} = T\mu$ and of $\Omega_C \equiv \operatorname{cov}\{v(C)\} = \operatorname{cov}\{v(V)\} = T\Omega_H T$ flow from those of μ and Ω_H using the spectral decomposition of T. Thus, in the important special case where $\mu = 0$ and $\Omega_H = \sigma^2 I_m$, one has $E(C) = B$, while the above-diagonal elements of C have covariance matrix

$$\Omega_C = (4n^2)^{-1}\sigma^2(P_J + 4P_W + nP_{HC}).$$

Explicit knowledge of Ω_C opens up the possibility of interval estimation for B and thereby the configuration $\tilde{X}$. Similar remarks apply to hypothesis testing.

The matrix R in the matrix representation of τ and κ also plays a role in the following probability model for H. Using the usual notation and assumptions, consider for $n > 3$ the following symmetric analysis of variance of model:

$$\text{for all} \quad i < j, \qquad h_{ij} = \theta + \alpha_i + \alpha_j + \epsilon_{ij},$$

in which we impose the usual identifying restriction that $\alpha \equiv (\alpha_1, \dots, \alpha_n)^T \in \mathbf{W}_n$. Using Proposition 4.1, this model can be written as

$$v(H) = \theta 1_m + R\alpha + \epsilon,$$

in which $\theta 1_m \in v(\mathbf{S}_I)$ and $R\alpha \in v(\mathbf{S}_W)$. Using the familiar characterization of the fitting process as orthogonal projection, we have from Propositions 4.5 and 4.6 that $\hat{\theta} 1_m = P_I v(H)$ and $R\hat{\alpha} = P_W v(H)$, so that

$$\hat{\theta} = n^{-1}(n-1)^{-1} h_{..} \quad \text{and} \quad \hat{\alpha}_i = (n-2)^{-1}\{h_i - n^{-1} h_{..}\}.$$

Using the orthogonal decomposition of H given in Proposition 4.6, we obtain at once the corresponding analysis of variance table:

Source	Sum of squares	Degrees of freedom
Mean	$\|H_I\|^2$	1
Between-point differences	$\|H_W\|^2$	$n-1$
Residual	$\|H_{HC}\|^2$	$m-n$
Total	$\|H\|^2$	m

A corresponding analysis is also possible for C.

Finally, we note a link with experimental design. We observe that R^T is the incidence matrix of a balanced incomplete block design with m blocks and n treatments, each block containing two distinct treatments, each treatment occurring in $n-1$ blocks, and each pair of distinct treatments occurring together in exactly one block. This design is symmetric if and only if $n = 3$.

It is a pleasure to gratefully acknowledge the value of the referee's comments on an earlier draft and the assistance provided by a Personal Research Grant from the British Social Science Research Council.

REFERENCES

Critchley, F. 1980. Optimal norm characterizations of multidimensional scaling methods and some related data analysis problems, in *Data Analysis and Informatics* (E. Diday, Ed.), North Holland, Amsterdam, pp. 209–229.

Critchley, F. 1986a. On certain linear transformations of the cone of symmetric nonnegative definite matrices, in preparation.

Critchley, F. 1986b. A new characterization of Euclidean distance matrices, in preparation.

Critchley, F. 1986c. A comparison of the classical and squared-distance methods of multidimensional scaling, in preparation.

Critchley, F. 1986d. Dimensionality theorems in multidimensional scaling and hierarchical cluster analysis, in *Data Analysis and Informatics IV* (E. Diday et al., Eds.), Elsevier Science, Amsterdam, pp. 45–70.

De Leeuw, J. and Heiser, W. 1982. Theory of multidimensional scaling, in *Handbook of Statistics, Volume 2* (P. R. Krishnaiah, Ed.), North-Holland, Amsterdam, Chapter 13.

Gower, J. C. 1982, Euclidean distance geometry, *Math. Sci.* 7:1–14.

Gower, J. C. 1984. Distance matrices and their Euclidean approximation, in *Data Analysis and Informatics III* (E. Diday, Ed.), North-Holland, Amsterdam, pp. 1–19.

Gower, J. C. 1985. Properties of Euclidean and non-Euclidean distance matrices, *Linear Algebra Appl.* 67:81–97.

Gower, J. C. 1986. Euclidean distance matrices, in *Multidimensional Data Analysis* (J. De Leeuw, Ed.), DSWO Press, Leiden Univ. to appear.

Holman, E. W. 1972. The relation between hierarchical clustering and Euclidean models for psychological distances, *Psychometrika* 37:417–423.

Kruskal, J. B. and Wish, M. 1978. *Multidimensional Scaling*, Sage, Beverly Hills, Calif.

Lingoes, J. C. 1971. Some boundary conditions for a monotone analysis of symmetric matrices, *Psychometrika* 36:195–203.

Luenberger, D. G. 1969. *Optimization by Vector Space Methods*, Wiley, New York.

Mathar, R. 1985. The best Euclidean fit to a given distance matrix in prescribed dimensions, *Linear Algebra Appl.* 67:1–6.

Rao, C. R. 1982. Diversity: Its measurement, decomposition, apportionment and analysis, *Sankhya Ser.* A 44:1–22.

Rao, C. R. 1984. Convexity properties of entropy functions and analysis of diversity, in *Inequalities in Statistics and Probability* (Y.L. Tong, Ed.), IMS Lecture Notes —Monograph Series, vol. 5.

Rao, C. F. 1986. Rao's axiomatization of diversity measures, in *Encyclopedia of Statistical Sciences* (Kotz and Johnson, Eds.), Wiley, New York, Vol. 7, pp. 614–617.

Torgerson, W. S. 1958. *Theory and Methods of Scaling*, Wiley, New York.

Received 1 April 1986; revised 11 August 1987.