Available online at www.sciencedirect.com
SCIENCE DIRECT®
LINEAR ALGEBRA AND ITS APPLICATIONS
ELSEVIER Linear Algebra and its Applications 397 (2005) 253–264
www.elsevier.com/locate/laa

On the diagonal scaling of Euclidean distance matrices to doubly stochastic matrices

Charles R. Johnson [a], Robert D. Masson [b], Michael W. Trosset [a,*]

[a] Department of Mathematics, College of William and Mary, P.O. Box 8795, Williamsburg, VA 23185-8795, USA
[b] Department of Mathematics, University of Chicago, 5734 S. University Avenue, Chicago, IL 60637, USA

Received 23 July 2001; accepted 25 October 2004

Submitted by U. Rothblum

Abstract

We consider the problem of scaling a nondegenerate predistance matrix A to a doubly stochastic matrix B. If A is nondegenerate, then there exists a unique positive diagonal matrix C such that $B = CAC$. We further demonstrate that, if A is a Euclidean distance matrix, then B is a spherical Euclidean distance matrix. Finally, we investigate how scaling a nondegenerate Euclidean distance matrix A to a doubly stochastic matrix transforms the points that generate A. We find that this transformation is equivalent to an inverse stereographic projection.

Keywords: Distance geometry; Stereographic projection

1. Preliminaries

A square matrix $B = (b_{ij})$ is *doubly stochastic* if and only if $b_{ij} \geqslant 0$ and $Be = e = B^{\mathrm{T}}e$, where $e = (1, \dots, 1)^{\mathrm{T}}$. A matrix A can be *scaled* to a doubly stochastic matrix B if and only if there exist strictly positive diagonal matrices D and E

* Corresponding author. Tel.: +1 757 221 2040; fax: +1 757 221 7400.
E-mail address: trosset@math.wm.edu (M.W. Trosset).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.10.023

such that $B = DAE$. The literature on scaling matrices to doubly stochastic matrices dates to at least as early as [11]. Only strictly positive A were considered in [11], but that hypothesis can be relaxed to obtain Lemma 1 below. Recall that A and C are *permutation equivalent* if there exist permutation matrices P and Q such $C = PAQ$, and that A is *completely irreducible* if it is irreducible via permutation equivalence.

Lemma 1. *If A can be scaled to a doubly stochastic matrix $B = DAE$, then B is the only doubly stochastic matrix to which A can be scaled. Furthermore, if A is completely irreducible, and if G and H are strictly positive diagonal matrices such that $B = GAH$, then there exists $t > 0$ such that $G = tD$ and $H = E/t$.*

The following result states which matrices can be scaled to doubly stochastic matrices. Derived from a theorem in [1], it is Remark 1 in [6].

Lemma 2. *A square nonnegative matrix A can be scaled to a doubly stochastic matrix if and only if A is permutation equivalent to a direct sum of completely irreducible matrices. In particular, A can be scaled to a doubly stochastic matrix if A is completely irreducible.*

We study the possibility of scaling *Euclidean distance matrices* to doubly stochastic matrices. The following terminology is becoming increasingly popular:

Definition 1. An $n \times n$ matrix $A = (a_{ij})$ is a Euclidean distance matrix (EDM) if and only if there exist $p_1, \ldots, p_n \in \Re^d$, $n \geqslant 2$ points in some d-dimensional Euclidean space, such that $a_{ij} = \|p_i - p_j\|^2$. The smallest d for which this is possible is the dimensionality of A.

Notice that the entries of an EDM are *squared* Euclidean distances, not the Euclidean distances themselves. It is evident from Definition 1 that, if $A = (a_{ij})$ is an EDM, then

1. $a_{ij} = \|p_i - p_j\|^2 \geqslant 0$ (A has nonnegative entries);
2. $a_{ij} = \|p_i - p_j\|^2 = \|p_j - p_i\|^2 = a_{ji}$ (A is symmetric); and
3. $a_{ii} = \|p_i - p_i\|^2 = 0$ (A is hollow).

We shall refer to a matrix that possesses these three properties as a *predistance matrix*. Furthermore, if each off-diagonal entry of the predistance matrix A is strictly positive, then we shall say that A is a *nondegenerate* predistance matrix. Notice that an EDM is nondegenerate if and only if it is generated by distinct points.

Although the sections that follow are not concerned with computation, we note that various researchers have proposed computational algorithms for the diagonal scaling of a nonnegative matrix A. A polynomial-time complexity bound on the problem of computing the scaling factors to a prescribed accuracy was derived in [8]. We do not know if the assumption that A is an EDM can be exploited to simplify computation. If A is an EDM, then a well-known constructive characterization of EDMs [10,14] allows one to compute a configuration of points that generate A. This calculation is the basis for classical multidimensional scaling [13,2] (not to be confused with diagonal scaling), a visualization technique that is popular in psychometrics and statistics.

2. Doubly stochastic scaling

We begin with a straightforward application of Lemma 2:

Theorem 1. *Every nondegenerate predistance matrix can be scaled to a doubly stochastic matrix.*

Proof. Let A be an $n \times n$ predistance matrix. If $n = 2$, then

$$A = \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}$$

with $a > 0$. Because

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

A is permutation equivalent to the direct sum of two completely irreducible matrices and the claim follows from Lemma 2.

Now suppose that $n > 2$. Let P and Q denote any two $n \times n$ permutation matrices. Because A is a nondegenerate predistance matrix, A has exactly one zero entry in each row and column. The matrix PA is a permutation of the rows of A, so it must have exactly one zero entry in each row and column. And the matrix PAQ is a permutation of the columns of PA, so it must have exactly one zero entry in each row and column. Therefore, it is impossible to find square matrices A_{11} and A_{22} that allow us to write PAQ in the form

$$PAQ = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}.$$

Thus, A is completely irreducible and again the claim follows from Lemma 2. □

We proceed to demonstrate that the scaling guaranteed by Theorem 1 can be written in a canonical form:

Theorem 2. *If A is a nondegenerate predistance matrix, then there exists a unique strictly positive diagonal matrix C such that CAC is doubly stochastic.*

Proof. By Theorem 1, there exist strictly positive diagonal matrices D and E such that $B = DAE$ is doubly stochastic. Because A is symmetric, $B^{\mathrm{T}} = EAD$. But B^{T}

is doubly stochastic because B is doubly stochastic; hence, it must be that $B^{\mathrm{T}} = B$ by virtue of Lemma 1.

We have shown that both DAE and EAD scale A to the same doubly stochastic matrix B. Now we suppose that $n \geqslant 3$, in which case A is completely irreducible and the second statement in Lemma 1 applies.

If A is completely irreducible, then there must exist $t > 0$ such that $E = tD$, in which case $B = tDAD$. Then $B = CAC$ upon setting $C = \sqrt{t}D$. This demonstrates existence, which can also be deduced by applying Corollary 2.2 in [7]. To demonstrate uniqueness, suppose that we had $B = CAC$ and $B = MAM$. Again applying Lemma 1, there must exist $t > 0$ such that $M = tC$ *and* $M = C/t$. Then $t = 1$ and $M = C$, as claimed.

The case $n = 2$ is covered by a straightforward calculation. □

Henceforth, whenever we refer to *the* scaling of a nondegenerate predistance matrix to a doubly stochastic matrix, we mean the scaling of Theorem 2.

Let A be a nondegenerate predistance matrix and let B denote the doubly stochastic matrix to which A scales. We have already remarked (in the proof of Theorem 2) that B is symmetric. Writing $B = CAC$ and $b_{ij} = c_{ii}a_{ij}c_{jj}$, we see that B is itself a nondegenerate predistance matrix. We now impose the additional assumption that A is an EDM and obtain the main result of this section:

Theorem 3. *Let A be a nondegenerate EDM and let B be the doubly stochastic matrix to which A scales. Then B is a nondegenerate EDM.*

Proof. Theorem 3.3 in [3] states that a nonzero EDM has exactly one positive eigenvalue. Now $B = CAC$ by Theorem 2, and Sylvester's Law of Inertia states that congruence relations preserve inertia, i.e., the numbers of positive and negative eigenvalues. (See, for example, Section 4.5 of [5].) It follows that B has exactly one positive eigenvalue. Furthermore, $Be = e$ because B is doubly stochastic. Combining these facts, it follows from Theorem 2.2 in [4] that B is an EDM. □

3. Doubly stochastic scaling and spherical distance matrices

Theorem 3 directs our attention to nondegenerate EDMs that are doubly stochastic. We proceed to investigate such matrices. Crucial to our investigation is the following:

Definition 2. An EDM is spherical if and only if it can be generated by points that lie on a sphere.

The following characterization of spherical distance matrices concatenates Theorem 3.4 in [12] and Theorem 2.2 in [3]. Recall that the centroid of a finite number of points in a vector space is the arithmetic mean of the points.

Lemma 3. *An $n \times n$ EDM B is spherical if and only if there exists $v \in \Re^n$ and $\lambda \geqslant 0$ such that $Bv = \lambda e$ and $v^{\mathrm{T}} e = 1$, in which case the radius of the sphere is $\sqrt{\lambda/2}$. Furthermore, the center of the sphere coincides with the centroid of the generating configuration if and only if e is an eigenvector of B.*

If the EDM B is doubly stochastic, then $Be = e$ and Lemma 3 applies with $v = e/n$ and $\lambda = 1/n$. Hence,

Theorem 4. *Let B denote an $n \times n$ EDM. Then B is doubly stochastic if and only if B is generated by a configuration of points that lie on a sphere whose center is the centroid of the configuration and whose radius is $\sqrt{1/2n}$.*

The fact that the doubly stochastic scaling of Theorem 2 scales arbitrary EDMs to spherical EDMs has several interesting consequences. For example, the following result was demonstrated in [3]:

Lemma 4. *Suppose that A is an EDM and* $\operatorname{rank}(A) = r$. *If A is spherical, then the dimensionality of A is $r - 1$; otherwise the dimensionality of A is $r - 2$.*

Because congruence relations preserve rank, it follows that $\operatorname{rank}(B) = \operatorname{rank}(CAC) = \operatorname{rank}(A)$. Hence,

Theorem 5. *Let A denote a nondegenerate EDM with dimensionality d. Let B denote the doubly stochastic EDM to which it scales. If A is spherical, then the dimensionality of B is d; otherwise the dimensionality of B is $d + 1$.*

A related consequence of Theorem 4 concerns the configurations of points that generate a nondegenerate EDM and the doubly stochastic EDM to which it scales. Somehow, doubly stochastic scaling transforms an arbitrary configuration into one that lies on a sphere. To investigate how, we performed several numerical experiments, described in [9]. Several features of the resulting configurations reminded David Lutzer of properties possessed by stereographic projection. In Section 4, we demonstrate that there is indeed an intimate connection between the doubly stochastic scaling of EDMs and stereographic projection.

4. Doubly stochastic scaling and stereographic projection

Given $r > 0$, let

$$S_d(r) = \left\{ x \in \Re^{d+1} : \sum_{i=1}^{d} x_i^2 + (x_{d+1} - r)^2 = r^2 \right\},$$

the sphere of radius r that is tangent to the hyperplane $x_{d+1} = 0$ at the origin $\theta = (0, \ldots, 0)^{\mathrm{T}}$ of $\Re^{d+1}$. The point $(0, \ldots, 0, 2r)^{\mathrm{T}} \in S_d(r)$ that is diametrically opposed to the point of tangency is called the north pole of $S_d(r)$. Given a point $p' \in S_d(r)$ that is not the north pole, let Λ denote the straight line that passes from the north pole through p'. The point p at which Λ intersects the hyperplane $x_{d+1} = 0$ is called the *stereographic projection* of p' into the hyperplane $x_{d+1} = 0$.

Stereographic projection defines a bijection between $S_d(r)$ less its north pole and the hyperplane $x_{d+1} = 0$. We will be interested in the mapping defined by the inverse of stereographic projection, which has the following explicit representation:

Lemma 5. *Let p denote the stereographic projection of $p' = (p'_1, \ldots, p'_d, p'_{d+1})^{\mathrm{T}} \in S_d(r)$ into the hyperplane $x_{d+1} = 0$. Then*

$$p'_{d+1} = \frac{2r\|p\|^2}{4r^2 + \|p\|^2} \quad \text{and} \quad p'_i = \frac{4r^2 p_i}{4r^2 + \|p\|^2} \tag{1}$$

for $i = 1, \ldots, d$.

Proof. Because Λ must pass through both $(0, \ldots, 0, 2r)^{\mathrm{T}}$, the north pole of $S_d(r)$, and $p = (p_1, \ldots, p_d, 0)^{\mathrm{T}}$, each point through which Λ passes can be written as

$$\Lambda(t) = (0, \ldots, 0, 2r)^{\mathrm{T}} + t(p_1, \ldots, p_d, -2r)^{\mathrm{T}}$$

for some $t \in \Re$.

We seek $t \neq 0$ for which $\Lambda(t) \in S_d(r)$. Let

$$t_p = \frac{4r^2}{4r^2 + \|p\|^2},$$

so that $\Lambda(t_p)$ is the point specified in (1). Because

$$\begin{aligned}\sum_{i=1}^{d} [t_p p_i]^2 + [(2r)(1 - t_p) - r]^2 &= t_p^2\|p\|^2 + r^2(1 - 2t_p)^2 \\ &= r^2 + t_p^2(4r^2 + \|p\|^2) - 4r^2 t_p = r^2,\end{aligned}$$

$\Lambda(t_p) \in S_d(r)$ and therefore $\Lambda(t_p) = p'$. □

Our argument that scaling an EDM to a doubly stochastic matrix is related to stereographic projection will rely on the following technical fact:

Lemma 6. *Let p and q denote the stereographic projections of $p', q' \in S_d(r)$ into the hyperplane $x_{d+1} = 0$. Then*

$$\|p' - q'\|^2 = \frac{16r^4\|p - q\|^2}{(4r^2 + \|p\|^2)(4r^2 + \|q\|^2)}.$$

Proof. Let $\rho = 4r^2$ and let

$$\kappa = \frac{\rho^2}{(\rho + \|p\|^2)^2(\rho + \|q\|^2)^2}.$$

Then, for $i = 1, \ldots, d$,

$$\begin{aligned}(p_i' - q_i')^2 &= \left(\frac{\rho p_i}{\rho + \|p\|^2} - \frac{\rho q_i}{\rho + \|q\|^2}\right)^2 \\ &= \kappa\big[(\rho + \|q\|^2)p_i - (\rho + \|p\|^2)q_i\big]^2 \\ &= \kappa\big[\rho^2(p_i - q_i)^2 + \|q\|^4 p_i^2 + \|p\|^4 q_i^2 + 2\rho\|q\|^2 p_i(p_i - q_i) \\ &\quad - 2\rho\|p\|^2 q_i(p_i - q_i) - 2\|p\|^2\|q\|^2 p_i q_i\big]\end{aligned}$$

and therefore

$$\begin{aligned}\frac{1}{\kappa}\sum_{i=1}^{d}(p_i' - q_i')^2 &= \rho^2\|p - q\|^2 + \|q\|^4\|p\|^2 + \|p\|^4\|q\|^2 \\ &\quad + 2\rho\big(\|q\|^2\|p\|^2 + \|p\|^2\|q\|^2\big) - 2\rho\big(\|p\|^2 + \|q\|^2\big)\langle p, q\rangle \\ &\quad - 2\|p\|^2\|q\|^2\langle p, q\rangle \\ &= \rho^2\|p - q\|^2 + \|p\|^2\|q\|^2\|p - q\|^2 + 4\rho\|p\|^2\|q\|^2 \\ &\quad - 2\rho\big(\|p\|^2 + \|q\|^2\big)\langle p, q\rangle.\end{aligned}$$

Also,

$$\begin{aligned}(p_{d+1}' - q_{d+1}')^2 &= \left(\frac{2r\|p\|^2}{\rho + \|p\|^2} - \frac{2r\|q\|^2}{\rho + \|q\|^2}\right)^2 \\ &= \frac{\rho}{\big(\rho + \|p\|^2\big)^2\big(\rho + \|q\|^2\big)^2}\big(\rho\|p\|^2 + \|p\|^2\|q\|^2 \\ &\quad - \rho\|q\|^2 - \|p\|^2\|q\|^2\big)^2 \\ &= \kappa\rho\big(\|p\|^4 - 2\|p\|^2\|q\|^2 + \|q\|^4\big);\end{aligned}$$

hence,

$$\begin{aligned}\frac{1}{\kappa}\|p' - q'\|^2 &= \frac{1}{\kappa}\sum_{i=1}^{d}(p_i' - q_i')^2 + \frac{1}{\kappa}(p_{d+1}' - q_{d+1}')^2 \\ &= \rho^2\|p - q\|^2 + \|p\|^2\|q\|^2\|p - q\|^2 + 4\rho\|p\|^2\|q\|^2 \\ &\quad - 2\rho\big(\|p\|^2 + \|q\|^2\big)\langle p, q\rangle + \rho\big(\|p\|^4 - 2\|p\|^2\|q\|^2 + \|q\|^4\big) \\ &= \rho^2\|p - q\|^2 + \|p\|^2\|q\|^2\|p - q\|^2 + \rho\big[4\|p\|^2\|q\|^2 \\ &\quad - 2\big(\|p\|^2 + \|q\|^2\big)\langle p, q\rangle + \|p\|^4 - 2\|p\|^2\|q\|^2 + \|q\|^4\big]\end{aligned}$$

$$
\begin{aligned}
&= \rho^2\|p-q\|^2 + \|p\|^2\|q\|^2\|p-q\|^2 + \rho\big[\big(\|p\|^2+\|q\|^2\big)^2 \\
&\quad - 2\big(\|p\|^2+\|q\|^2\big)\langle p,q\rangle\big] \\
&= \rho^2\|p-q\|^2 + \|p\|^2\|q\|^2\|p-q\|^2 \\
&\quad + \rho\big(\|p\|^2+\|q\|^2\big)\|p-q\|^2 \\
&= \big(\rho+\|p\|^2\big)\big(\rho+\|q\|^2\big)\|p-q\|^2.
\end{aligned}
$$

Multiplying both sides of this expression by κ now yields the desired result. □

An immediate consequence of Lemma 6 is that stereographic projection is related to the diagonal scaling of EDMs.

Theorem 6. *Let $\{p_1,\ldots,p_n\}$ denote the stereographic projections of $p'_1,\ldots,p'_n \in S_d(r)$ into the hyperplane $x_{d+1}=0$. Let A and A' denote the EDMs that correspond to these configurations and let D denote the diagonal matrix with diagonal entries*

$$d_{ii} = \frac{4r^2}{4r^2+\|p_i\|^2}.$$

Then $A' = DAD$.

Proof. Applying Lemma 6,

$$
\begin{aligned}
a'_{ij} &= \|p'_i - p'_j\|^2 \\
&= \frac{16r^4\|p_i-p_j\|^2}{\big(4r^2+\|p_i\|^2\big)\big(4r^2+\|p_j\|^2\big)} \\
&= \frac{4r^2}{\big(4r^2+\|p_i\|^2\big)} a_{ij} \frac{4r^2}{\big(4r^2+\|p_j\|^2\big)} = d_{ii}a_{ij}d_{jj}. \qquad \square
\end{aligned}
$$

We now begin a somewhat intricate argument that will culminate in our main result. To make this argument, it is convenient to consider stereographic projection with respect to

$$S_d(z,r) = \left\{ x \in \Re^{d+1} : \sum_{i=1}^{d}(x_i-z_i)^2 + (x_{d+1}-r)^2 = r^2 \right\},$$

the sphere of radius r that is tangent to the hyperplane $x_{d+1}=0$ at $(z^{\mathrm T},0)^{\mathrm T} \in \Re^{d+1}$. Notice that the inverse stereographic projection of a point in $x_{d+1}=0$ that is far from $(z^{\mathrm T},0)^{\mathrm T}$ will be near the north pole of $S_d(z,r)$. This observation is quantified by the following inequality:

Lemma 7. *Let p denote the stereographic projection with respect to $S_d(z,r)$ of p' into the hyperplane $x_{d+1}=0$. If $0 < r \leqslant \|p-z\|/\big(2\sqrt{3}\big)$, then $p'_{d+1} \geqslant 3r/2$.*

Proof. By choosing a coordinate system in which $(z^{\mathrm{T}}, 0)^{\mathrm{T}}$ is the origin of $\Re^{d+1}$, we can use Lemma 5 to calculate that

$$p'_{d+1} = \frac{2r\|p-z\|^2}{4r^2 + \|p-z\|^2} \geqslant \frac{2r\|p-z\|^2}{\|p-z\|^2/3 + \|p-z\|^2} = \frac{3r}{2}. \qquad \square$$

Now let P denote a configuration of points $p_1, \ldots, p_n \in \{x \in \Re^{d+1} : x_{d+1} = 0\}$. Given the sphere $S_d(z, r)$, let $P(z, r)$ denote the configuration of points obtained by inverse stereographic projection, i.e., the configuration of points $p'_1, \ldots, p'_n \in \Re^{d+1}$ such that p_i is the stereographic projection with respect to $S_d(z, r)$ of p'_i into the hyperplane $x_{d+1} = 0$. We focus on which sphere is used for stereographic projection. The following result is crucial to our investigation:

Lemma 8. *Let P denote a configuration of $n \geqslant 2$ distinct points in $\{x \in \Re^{d+1} : x_{d+1} = 0\}$. There exists a sphere $S_d(z, r)$ whose center is the centroid of the configuration $P(z, r)$.*

Proof. We give separate arguments for $n = 2$ and $n \geqslant 3$. If $n = 2$, then let z denote the centroid of P in $x_{d+1} = 0$ and choose a coordinate system in which $(z^{\mathrm{T}}, 0)^{\mathrm{T}}$ is the origin of $\Re^{d+1}$, so that $P = \{p, -p\}$ and the center of $S_d(z, r)$ is $(0, \ldots, 0, r)^{\mathrm{T}}$. Then it follows from Lemma 5 that the centroid of $P(z, r)$ has coordinates

$$\frac{1}{2}\left[\frac{4r^2 p_i}{4r^2 + \|p\|^2} + \frac{4r^2(-p_i)}{4r^2 + \|p\|^2}\right] = 0$$

for $i = 1, \ldots, d$ and

$$\frac{1}{2}\left[\frac{2r\|p\|^2}{4r^2 + \|p\|^2} + \frac{2r\|-p\|^2}{4r^2 + \|p\|^2}\right] = \frac{2r\|p\|^2}{4r^2 + \|p\|^2},$$

which equals r if $r = \|p\|/2$.

Now suppose that $n \geqslant 3$ and let K be any bounded convex subset of $\{x \in \Re^{d+1} : x_{d+1} = 0\}$ that contains P. Let $f : K \times (0, \infty) \to \Re^{d+1}$ denote the function that maps (z, r) to the centroid of $P' = P(z, r)$. We will show that f has a fixed point.

The continuity of f follows from an application of Lemma 5 to $P - z$. Furthermore, the $(d + 1)$st coordinate function of f is

$$h(z, r) = \frac{1}{n}\sum_{i=1}^{n} \frac{2r\|p_i - z\|^2}{4r^2 + \|p_i - z\|^2}.$$

Notice that h is strictly positive on $K \times (0, \infty)$.

Next we establish that h is bounded above. Let $\beta < \infty$ denote the diameter of K, finite because K is bounded. Let $\delta_i = \|p_i - z\| \leqslant \beta$. Then

$$h(z, r) \leqslant \frac{1}{n}\sum_{i=1}^{n} \delta_i \frac{2r\delta_i}{r^2 + \delta_i^2} \leqslant \frac{\beta}{n}\sum_{i=1}^{n} \frac{2r\delta_i}{(r - \delta_i)^2 + 2r\delta_i} \leqslant \beta < \infty.$$

Next we investigate the behavior of h when r is small. Let

$$\epsilon = \frac{1}{2}\min_{i \neq j} \|p_i - p_j\|,$$

strictly positive because we have assumed that the points in P are distinct. For any $z \in K$, there can be at most one point $p \in P$ for which $\|p - z\| < \epsilon$. Let $\gamma = \epsilon/\big(2\sqrt{3}\big)$ and suppose that $r < \gamma$. By Lemma 7, if $\|p - z\| \geqslant \epsilon$, then the $(d+1)$st coordinate of p' is at least $3r/2$. Because $n \geqslant 3$,

$$h(z,r) \geqslant \frac{1}{n}\sum_{i=1}^{n-1}\frac{3r}{2} = \frac{n-1}{n} \times \frac{3r}{2} \geqslant r. \tag{2}$$

Now let α denote the minimum of the continuous function h on the compact set $K \times [\gamma, \beta]$. Because h is strictly positive on $K \times (0, \infty)$, $\alpha > 0$. Let $L = K \times [\alpha, \beta]$. Evidently, L is compact and convex. We claim that f maps L into itself.

First, given $(z^{\mathrm{T}}, r) \in L$, we choose a coordinate system in which $(z^{\mathrm{T}}, 0)^{\mathrm{T}}$ is the origin. Suppose that $x \in K$. Then, by Lemma 5,

$$x_i' = \left(\frac{4r^2}{4r^2 + \|x\|^2}\right) x_i,$$

so $(x_1', \ldots, x_d')^{\mathrm{T}}$ lies on the line segment that connects x and z. Because K is convex, $x' \in K \times (0, \infty)$. It follows that, if $P \subset K$, then the centroid of P' lies in $K \times (0, \infty)$.

Second, we consider the consequences of choosing $r \in [\alpha, \beta]$. By the definition of β, $h(z, r) \leqslant \beta$. If $r \geqslant \gamma$, then $h(z, r) \geqslant \alpha$ by the definition of α; if $r < \gamma$, then $h(z, r) \geqslant r \geqslant \alpha$ by (2).

We conclude that f maps the compact and convex set L into itself. The existence of a fixed point then follows from the Brouwer fixed point theorem. □

We can now state the first of our two main results.

Theorem 7. *Let $P \subset \Re^d$ denote a configuration of at least two distinct points. Let A denote the EDM that corresponds to P and let B denote the doubly stochastic matrix to which A scales. There exists (z, r) such that A', the EDM that corresponds to $P' = P(z, r)$, is a scalar multiple of B.*

Proof. Invoking Lemma 8, let $S_d(z, r)$ be a sphere whose center is the centroid of $P' = P(z, r)$. Then A' is a spherical EDM and it follows from Lemma 3 that e is an eigenvector of A'. Hence, A' is a scalar multiple of a doubly stochastic matrix, say $A' = \sigma B'$.

By Theorem 6, $A' = DAD$ for a strictly positive diagonal matrix D. Setting $C = D/\sqrt{\sigma}$, we see that $CAC = DAD/\sigma = A'/\sigma = B'$, i.e., A scales to the doubly stochastic matrix B'. By Lemma 1, $B' = B$. □

Theorem 7 expresses a relation between an EDM and stereographic projection in terms of the EDM. Our final result expresses the same relation in terms of the configuration of points that generates the EDM.

Theorem 8. *Let $P \subset \Re^d$ denote a configuration of $n \geqslant 2$ distinct points. Let A denote the EDM that corresponds to P and let B denote the doubly stochastic matrix to which A scales. There exists an affine linear transformation $T : \Re^n \to \Re^n$ such that B is the EDM that corresponds to the configuration Q', the inverse stereographic projection with respect to $S_d\left(1/\sqrt{2n}\right)$ of the transformed configuration $Q = T(P)$.*

Proof. Invoking Lemma 8, let $S_d(z, r)$ be a sphere whose center is the centroid of $P' = P(z, r)$. Let $t = r\sqrt{2n}$ and define $T : \Re^n \to \Re^n$ by $T(x) = (x - z)/t$. Because T is a translation followed by a dilation, the EDM that corresponds to the configuration $Q = T(P)$ is A/t^2.

Let Q' denote the inverse stereographic projection with respect to $S_d\left(1/\sqrt{2n}\right)$ of Q. We choose a coordinate system whose origin is z and apply Theorem 6 to conclude that $A' = D(A/t^2)D = CAC$, where $C = D/t$. Thus, A can be diagonally scaled to A'.

Because the centroid of P' is the center of $S_d(z, r)$, the centroid of $[P - z]'$ is the center of $S_d(r)$ and therefore the centroid of $Q' = [(P - z)/t]'$ is the center of $S_d(r/t) = S_d\left(1/\sqrt{2n}\right)$. Because $Q' \subset S_d\left(1/\sqrt{2n}\right)$, it follows from Theorem 4 that A' is doubly stochastic. By Lemma 1, the doubly stochastic matrix to which A scales is unique; hence, $B = A'$. $\square$

Acknowledgments

This research was inspired by an inquiry from John Crow. It was conducted in the summer of 2000, when the second author participated in the REU program *Matrix Analysis and Its Applications* at the College of William and Mary in Williamsburg, VA, and was mentored by the other two authors. The REU program was funded by grant DMS-99-87803 from the National Science Foundation. We are grateful to David Lutzer for his dedicated administration of that program, for his keen interest in this research, and for his valuable contributions to it. We are also grateful to six anonymous referees, whose suggestions improved the quality of this manuscript.

References

[1] R. Brualdi, S. Parter, H. Schneider, The diagonal equivalence of a nonnegative matrix to a stochastic matrix, J. Math. Anal. Appl. 16 (1966) 31–50.

[2] J.C. Gower, Some distance properties of latent root and vector methods in multivariate analysis, Biometrika 53 (1966) 315–328.

[3] T.L. Hayden, P. Tarazaga, Distance matrices and regular figures, Linear Algebra Appl. 195 (1993) 9–16.
[4] T.L. Hayden, R. Reams, J. Wells, Methods for constructing distance matrices and the inverse eigenvalue problem, Linear Algebra Appl. 295 (1999) 97–112.
[5] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[6] C.R. Johnson, R.B. Kellogg, An inequality for doubly stochastic matrices, J. Res. National Bureau Standards B 80 (1976) 433–436.
[7] B. Kalantari, A theorem of the alternative for multihomogeneous functions and its relationship to diagonal scaling of matrices, Linear Algebra Appl. 236 (1996) 1–24.
[8] B. Kalantari, L. Khachiyan, On the complexity of nonnegative-matrix scaling, Linear Algebra Appl. 240 (1996) 87–103.
[9] R.D. Masson, On the diagonal scaling of squared distance matrices to doubly stochastic matrices, REU Report, Department of Mathematics, College of William and Mary, Williamsburg, VA, 2000.
[10] I.J. Schoenberg, Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espaces distanciés vectoriellement applicable sur l'espace de Hilbert", Ann. Math. 38 (1935) 724–732.
[11] R. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices, Ann. Math. Stat. 35 (1964) 876–879.
[12] P. Tarazaga, T.L. Hayden, J. Wells, Circum–Euclidean distance matrices and faces, Linear Algebra Appl. 232 (1996) 77–96.
[13] W.S. Torgerson, Multidimensional scaling: I. Theory and method, Psychometrika 17 (1952) 401–419.
[14] G. Young, A.S. Householder, Discussion of a set of points in terms of their mutual distances, Psychometrika 3 (1938) 19–22.