SCALING RADIAL BASIS FUNCTIONS VIA EUCLIDEAN DISTANCE MATRICES

B. J. C. BAXTER

ABSTRACT. A radial basis function approximation is typically a linear combination of shifts of a radially symmetric function, possibly augmented by a polynomial of suitable degree, that is, it takes the form

$$s(x) = \sum_{k=1}^{n} c_k \phi(\|x - x_k\|) + p(x), \qquad x \in \mathbb{R}^d.$$

In the mid 1980s, Micchelli, building on pioneering work of Schoenberg in the 1930s and 1940s, provided simple sufficient conditions on ϕ that imply radial basis functions can interpolate scattered data. However, when the data density varies locally, several authors, such as Hon and Kansa [5], have suggested scaling the translates. In other words, it can be advantageous to replace the Euclidean norm by some more general distance functional $\Delta(\cdot,\cdot)$, that is

$$s(x) = \sum_{k=1}^{n} c_k \phi(\Delta(x, x_k)) + p(x), \qquad x \in \mathbb{R}^d.$$

This distance functional Δ need not be a metric, but we shall require that Δ be symmetric and satisfy $\Delta(x,x) = 0$, for all $x \in \mathbb{R}^d$. Unfortunately, the Micchelli-Schoenberg theory does not obviously apply in this more general setting, but some papers have observed that interpolation is well-defined if the distance functional is a sufficiently small perturbation of the Euclidean norm. However, in this study we follow a different approach which returns to the roots of Schoenberg's work. Specifically, we use Schoenberg's classification of Euclidean distance matrices to provide a simple technique which, given a suggested distance functional Δ, calculates a perturbed distance functional $\widehat{\Delta}$ for which the underlying interpolation matrix is invertible, when the function ϕ is strictly positive definite (i.e. a Mercer kernel) or strictly conditionally positive (or negative) definite of order one. As a simple by-product of this method, we can also apply the Narcowich-Ward [10] norm estimate results easily, since the minimum distance between points is now under our control via $\widehat{\Delta}$.

1. Introduction

A radial basis function is typically an approximation of the form

$$s(x) = \sum_{k=1}^{n} c_k \phi(\|x - x_k\|) + p(x), \qquad x \in \mathbb{R}^d, \tag{1.1}$$

where $\phi : [0,\infty) \to \mathbb{R}$, $c_1, \ldots, c_n$ are real numbers, p is a polynomial, and $\|\cdot\|$ denotes the Euclidean norm. Such functions have proved themselves to be of great practical and theoretical importance since C. A. Micchelli established rather mild conditions under which they provide interpolants to multivariate scattered data; see, for instance, the useful book of Buhmann [7], and, of course, Micchelli's seminal paper [9]. Specifically, let m be a non-negative integer and let $p_1, \ldots, p_M$ be any basis for the M-dimensional vector space $P_m(\mathbb{R}^d)$ of polynomials of degree m on $\mathbb{R}^d$. Micchelli derived conditions for which the *augmented interpolation matrix*

$$A_P = \begin{pmatrix} A & P \\ P^T & 0 \end{pmatrix} \in \mathbb{R}^{(n+M)\times(n+M)} \tag{1.2}$$

is invertible, where

$$A_{jk} = \phi(\|x_j - x_k\|), \qquad 1 \le j, k \le n, \tag{1.3}$$

and

$$P_{jk} = p_k(x_j), \qquad 1 \le j \le n, \quad 1 \le k \le N. \tag{1.4}$$

Thus a necessary condition for nonsingularity of A_P is that the matrix P be of full rank, which imposes a geometric constraint on the points $x_1, \ldots, x_n$: if P is of full rank, then we say that the data $x_1, \ldots, x_n$ are $P_m(\mathbb{R}^d)$ *unisolvent*; in other words, no nontrivial polynomial of degree m can vanish at every data point $x_1, \ldots, x_n$. The sufficient condition is that $y^T A y \ge 0$ when $P^T y = 0$, with equality if and only if $y = 0$; we say that ϕ is *strictly conditionally positive definite of order* m if this property obtains.

We now require some of the details of Micchelli's analysis. To this end, we remind the reader that a function $f : [0, \infty) \to \mathbb{R}$ is *completely monotonic* if it is infinitely differentiable and satisfies $(-1)^k f^{(k)}(t) \ge 0$, for all $t > 0$ and any non-negative integer k. [The celebrated Bernstein–Hausdorff–Widder theorem characterizes completely monotonic functions as Laplace transforms of positive measures on the half-line $[0, \infty)$.]

Theorem 1.1. *Let m be a non-negative integer, let $f : [0, \infty) \to \mathbb{R}$ be any function for which $(-1)^m f^{(m)}$ is a nonconstant completely monotonic function, and define $\phi(r) = f(r^2)$, for $r \ge 0$. Then, for any positive integers n and d, the augmented interpolation matrix A_P defined by* (1.2) *is invertible if the points $x_1, \ldots, x_n$ form a $P_m(\mathbb{R}^d)$ unisolvent set.*

Proof. This is Theorem 2.1 in [9]. □

The power of this theorem lies in the beautiful fact that we can use ϕ to interpolate in any ambient dimension d, and that it is easy to construct such functions. For example, it is easily checked that $f(t) = (t + c^2)^{-1/2}$ is completely monotonic, so that $\phi(r) = (r^2 + c^2)^{-1/2}$ is the inverse multiquadric and Theorem 1.1 implies that the matrix A defined by

$$A_{jk} = f(\|x_j - x_k\|^2), \qquad 1 \le j, k \le n,$$

is positive definite (a Mercer kernel, to use the historically precise terminology of Learning Theory) for any distinct points $x_1, \ldots, x_n$ lying in any $\mathbb{R}^d$ – we say that f is a positive definite function on Hilbert space. The author has recently provided [2] a geometric proof of the fact that positive definite functions on Hilbert space are completely monotonic.

As a second illustration of Theorem 1.1, we let $f(t) = -(t + c^2)^{1/2}$, so that $-f'(t) = (1/2)(t + c^2)^{-1/2}$ and we deduce that the Hardy multiquadric $\phi(r) = (r^2 + c^2)^{1/2}$ is strictly conditionally negative definite of order one.

It is not obvious that Theorem 1.1 tells us anything when our norm is not the Euclidean norm. However, this is not so. For example, the author extended Micchelli's results as follows.

Theorem 1.2. *Let the norm* $\|.\|$ *of* (1.2) *be replaced by any* p*-norm, for* $1 \le p \le 2$, *that is*

$$\|x\|_p = \left(\sum_{k=1}^{d} |x_k|^p \right)^{1/p}, \qquad x \in \mathbb{R}^d.$$

Then the p*-norm interpolation matrix* $A \in \mathbb{R}^{n \times n}$*, defined by*

$$A_{jk} = \|x_j - x_k\|_p, \qquad 1 \le j, k \le n,$$

is nonsingular when $n \ge 2$ *and the points* $x_1, \ldots, x_n$ *are distinct. Further* $y^T A y \le 0$ *when the components of the vector* $y \in \mathbb{R}^n$ *sum to zero, with equality if and only if* y *is the zero vector.*

Proof. This is Theorem 2.11 of [1]. □

In other words, we can replace radial basis functions by *p-norm radial basis functions*, that is we can interpolate scattered multivariate data using functions of the form

$$s(x) = \sum_{k=1}^{n} c_k \phi(\|x - x_k\|_p), \qquad x \in \mathbb{R}^d,$$

when the conditions of Theorem 1.2 are satisfied. [The reader might be intrigued to learn that we *cannot* use p-norms for $p > 2$; see Section 3 of [1].] Now, this paper is certainly *not* suggesting that p-norm radial basis functions provide useful practical alternatives to radial basis functions (although no numerical experiments have been published, to the author's knowledge). Instead, the kernel of the proof of Theorem 1.2 suggested the idea of the present paper. Specifically, the author demonstrated in [1] that the p-norm distance matrix was a *Euclidean distance matrix*, that is, there exist vectors $z_1, \ldots, z_n \in \mathbb{R}^n$ for which

$$\|x_j - x_k\|_p = \|z_j - z_k\|^2, \qquad 1 \le j, k \le n,$$

and details are provided in the next section. Once this is established, we can shift attention to the vectors $z_1, \ldots, z_n \in \mathbb{R}^n$ and apply Theorem 1.1, because the change in ambient dimension from d to n is permissible. This suggests the possibility that some alternative distance functionals can be used. Indeed, if we are sufficiently fortunate that the *distance functional matrix* $D \in \mathbb{R}^{n \times n}$, defined by

$$D_{jk} = \Delta(x_j, x_k), \qquad 1 \le j, k \le n,$$

is a Euclidean distance matrix, then Theorem 1.1 implies the invertibility of the augmented interpolation matrix. Of course, it is rather unlikely that D will be a Euclidean distance matrix, so it is natural to consider methods for perturbing the distance functional. We describe a simple technique which provides a new functional $\widehat{\Delta}$ for which the corresponding distance functional matrix $\widehat{D}$ is a Euclidean distance matrix.

Theorem 1.3. *Let* $\Delta : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *be any symmetric distance function, that is,* $\Delta(x, y) = \Delta(y, x)$*, for all* $x, y \in \mathbb{R}^d$*, which satisfies* $\Delta(x, x) = 0$*, for all* $x \in \mathbb{R}^d$*. Let* μ *be any positive constant exceeding*

$$\max_{1 \le j \le n-1} \left(-2D_{jn} + \sum_{k=1, k \ne j}^{n} |D_{jn} + D_{kn} - D_{jk}| \right).$$

Then the distance functional $\widehat{\Delta} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ *defined by*

$$\widehat{\Delta}(x, x_n) = \Delta(x, x_n) + \mu/2, \qquad x \in \mathbb{R}^d \setminus \{x_n\},$$

and

$$\widehat{\Delta}(x, y) = \Delta(x, y) + \mu,$$

for any distinct points $x, y \in \mathbb{R}^d \setminus \{x_n\}$*, generates a new distance functional matrix* $\widehat{D}$ *which is a Euclidean distance matrix*

This theorem is proved in the following section. However, note that the point x_n could be replaced by any one of $x_1, \ldots, x_{n-1}$, although it is likely that the numerical properties of the method are sensitive to the choice of x_n.

Given Theorem 1.3, we obtain a simple interpolation result for positive definite functions.

Theorem 1.4. *Let* $f : [0, \infty) \to \mathbb{R}$ *be any non-constant completely monotonic functions and define* $\phi(r) = f(r^2)$*, for* $r \geq 0$*. If the distance functional* Δ *is perturbed to form* $\widehat{\Delta}$*, as described in Theorem 1.3, then the unaugmented interpolation matrix* A *defined by* (1.3) *is invertible.*

Proof. The construction of the modified distance functional $\widehat{\Delta}$ implies that

$$\widehat{\Delta}(x_j, x_k) = \|z_j - z_k\|^2 > 0, \qquad \text{when } j \neq k, \tag{1.5}$$

so that Theorem 1.1 implies the invertibility of (1.3). □

This is a new technique and many points remain for future study. Nevertheless, it enables us to construct nonsingular interpolation matrices in a rather simple way. Further, the minimum distance between the vectors $z_1, \ldots, z_n$ can be used in the Narcowich–Ward invertibility theorems [10] , if needed. However, the extension to conditionally positive definite functions of higher orders remains unclear. For example, if ϕ were a strictly conditionally negative definite function of order m, for $m > 1$, then the new points $z_1, \ldots, z_n$ defined by (1.5) might no longer be unisolvent, although this seems unlikely. Fortunately, this problem does not occur in the important special case of conditionally negative definite functions of order one.

Theorem 1.5. *Let* $f : [0, \infty) \to \mathbb{R}$ *be any function for which* f' *is a nonconstant completely monotonic function and* $f(0) \geq 0$*. If the distance functional* Δ *is perturbed to form* $\widehat{\Delta}$*, as described in Theorem 1.3, then the unaugmented interpolation matrix* A *defined by* (1.3) *and the augmented interpolation matrix* A_P *defined by* (1.2) *are both invertible.*

Proof. The construction of the modified distance functional $\widehat{\Delta}$ implies that

$$\widehat{\Delta}(x_j, x_k) = \|z_j - z_k\|^2 > 0, \qquad \text{when } j \neq k. \tag{1.6}$$

Further the locations of the points $z_1, \ldots, z_n \in \mathbb{R}^n$ play no part in the definition of $P_1(\mathbb{R}^n)$-unisolvent, which simply requires that the vector $y \in \mathbb{R}^n$ has components summing to zero. Thus the augmented interpolation matrix A_P is invertible. The unaugmented interpolation matrix A satisfies $y^T A y < 0$, when $y \in \mathbb{R}^n$ is any nonzero vector whose components sum to zero, and the set of such zero-summing vectors forms a vector space of dimension $n-1$. Hence, following Micchelli [9], A must have at least $n-1$ negative eigenvalues. Since $f(0) \geq 0$, the trace of the matrix A, which is the sum of its eigenvalues, must be non-negative, so the remaining eigenvalue is positive. Hence A is invertible. $\square$

One significant disadvantage of the perturbed distance functional $\widehat{\Delta}$ is that it is not a continuous function. Therefore the author is considering alternative modifications that preserve continuity, and intends to report on these soon.

2. Euclidean Distance Matrices

The classical theory of radial basis function interpolation is a series of footnotes to Schoenberg's brilliant analysis of the isometric embedding problems, evolved in a series of papers [11, 12, 13]. These have been collected in [14], edited by Carl de Boor, together with highly useful commentary. We now delve into Schoenberg's rich legacy.

Definition 2.1. *We shall say that a matrix $M \in \mathbb{R}^{n\times n}$ is a* Euclidean distance matrix *if there exist vectors $z_1, \ldots z_n \in \mathbb{R}^n$ for which*

$$M_{jk} = \|z_j - z_k\|^2, \qquad 1 \le j,k \le n.$$

We let E_n denote the set of all $n\times n$ Euclidean distance matrices.

It transpires that $n \times n$ Euclidean distance matrices are really $(n-1)\times(n-1)$ non-negative definite symmetric matrices in disguise, as revealed by an important characterization theorem due to Schoenberg, which is our next topic. Now, it is clearly necessary that a Euclidean distance matrix must be symmetric and have zero diagonal elements, and this larger class of matrices will be useful in its own right.

Definition 2.2. *We shall say that a symmetric matrix $D \in \mathbb{R}^{n\times n}$ is* almost Euclidean *if D is a symmetric matrix whose diagonal elements vanish. The set of $n \times n$ almost Euclidean matrices will be denoted by $Æ_n$.*

Let Symm_k denote the linear space of symmetric matrices in $\mathbb{R}^{k\times k}$ and let P_k denote the convex cone of non-negative definite matrices in Symm_k. We can now state the fundamental geometric characterization of Schoenberg, providing proofs in the modern idiom for the convenience of the reader.

Theorem 2.3. *Define $\tau : Æ_n \to \mathrm{Symm}_{n-1}$ by*

$$\tau(A)_{jk} = A_{jn} + A_{kn} - A_{jk}, \qquad 1 \le j,k \le n-1. \tag{2.1}$$

Then τ is a linear bijection between $Æ_n$ and Symm_{n-1}. Further, given $M \in \mathrm{Symm}_{n-1}$, we have

$$\tau^{-1}(M)_{jk} = \begin{cases} \frac{1}{2}M_{jj}, & 1 \le j \le n-1, k = n \\ \frac{1}{2}M_{kk}, & j = n, 1 \le k \le n-1 \\ \frac{1}{2}\left(M_{jj} + M_{kk} - 2M_{jk}\right), & 1 \le j,k \le n-1, \\ 0 & j = k = n. \end{cases} \tag{2.2}$$

Proof. The map τ is clearly linear. Since the dimensions of $Æ_n$ and Symm_{n-1} are both equal to $1 + 2 + \cdots + n - 1$, we need only prove that τ is injective. To this end, suppose $\tau(A) = 0$ and set $j = k$ in (2.1). Since the diagonal elements almost Euclidean matrices vanish, by definition, we deduce that $A_{jn} = 0$, for $j =$

$1, \ldots, n-1$. If we now choose $j \neq k$ in (2.1), we see that $A_{jk} = 0$. Hence τ is a linear bijection. An almost identical calculation yields (2.2). $\square$

Theorem 2.4. *Let $A \in Æ_n$. Then $A \in E_n$ if and only if $\tau(A) \in P_{n-1}$.*

Proof. If $A \in E_n$, that is,

$$A_{jk} = \|x_j - x_k\|^2, \qquad 1 \le j, k \le n.$$

We may shift the points $x_1, \ldots, x_n$ without changing their mutual distances, so we can, and do, assume $x_n = 0$. Thus

$$\tau(A)_{jk} = \|x_j\|^2 + \|x_k\|^2 - \|x_j - x_k\|^2 = 2x_j^T x_k, \qquad 1 \le j, k \le n-1,$$

and thus we have shown that $\tau(A)$ is non-negative definite, being a Gram matrix.

Conversely, if $M = \tau(A) \in P_{n-1}$, then we can write

$$M_{jk} = 2v_j^T v_k, \qquad 1 \le j, k \le n-1.$$

Hence, if we define $v_n = 0$, then (2.2) implies

$$\tau^{-1}(M)_{jk} = \|v_j - v_k\|^2, \qquad 1 \le j, k \le n.$$

$\square$

As a simple corollary, we mention the calculation of $x_1, \ldots, x_n$ given $A \in E_n$.

Corollary 2.5. *Let $A \in E_n$. Let $r_1, \ldots, r_{n-1} \in \mathbb{R}^n$ be any vectors generating the matrix $\tau(A) \in P_{n-1}$, that is,*

$$r_j^T r_k = \tau(A)_{jk}, \qquad 1 \le j, k \le n-1.$$

Then $A_{jk} = \|x_j - x_k\|^2$, for $1 \le j, k \le n-1$, where $x_n = 0$ and $x_j = r_j/\sqrt{2}$, for $1 \le j \le n-1$.

For example, if we compute the Cholesky factorization $\tau(A) = R^T R$, where $R \in \mathbb{R}^{(n-1)\times(n-1)}$, then we simply take each column of R, embed it in $\mathbb{R}^n$ by defining its nth component to be zero, and apply the theorem. We also observe that the rank of R is the minimal dimension of Euclidean space in which the simplex formed by $x_1, \ldots, x_n$ can be embedded.

We have already mentioned that E_n is a convex cone, and this follows from the relation $E_n = \tau^{-1}(P_{n-1})$. If we now use the inner product on $\mathbb{R}^{n\times n}$ matrices induced by the Frobenius norm, that is

$$\langle C, D\rangle_F = \sum_{k=1}^{n}\sum_{\ell=1}^{n} C_{kl} D_{kl} \tag{2.3}$$

and

$$\|C\|_F = \left(\sum_{k=1}^{n}\sum_{\ell=1}^{n} C_{kl}^2\right)^{1/2}, \tag{2.4}$$

then a well-known classical result of Hilbert space theory implies that, given any element $D \in Æ_n$, there exists a unique closest Euclidean distance matrix $\hat{D} \in E_n$. There are some applications for which it is important to find $\hat{D}$; see, for instance Higham [4], Gower [3]. However, we are not required to use the Frobenius norm on $Æ_n$. Instead, we can let our inner product on $Æ_n$ be the so-called pull-back inner product. Specifically, we may define

$$\langle D_1, D_2\rangle = \langle \tau(D_1), \tau(D_2)\rangle_F, \qquad \text{for } D_1, D_2 \in Æ_n, \tag{2.5}$$

so that

$$\|D_1 - D_2\| = \|\tau(D_1) - \tau(D_2)\|_F, \qquad \text{for } D_1, D_2 \in \text{Æ}_n. \tag{2.6}$$

This was suggested by Gower [3] and Mathar [8], because of the excellent reason that matrix nearness problems involving Symm_{n-1} are well-understood, and is of some importance in computational chemistry and multivariate data analysis (see, for instance, [6]). Specifically, given the spectral decomposition

$$M = Q_M \Lambda_M Q_M^T, \qquad M \in \mathrm{Symm}_{n-1},$$

where Q_M is orthogonal and $\Lambda_M = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix formed by the eigenvalues of M, the closest element $\widehat{M} \in P_{n-1}$ is given by

$$\widehat{M} = Q_M \widehat{\Lambda_M} Q_M^T,$$

where

$$\widehat{\Lambda_M} = \mathrm{diag}\left(\max\{\lambda_1, 0\}, \ldots, \max\{\lambda_n, 0\}\right).$$

Thus, given any almost Euclidean distance matrix $A \in \text{Æ}_n$, we calculate $\tau^{-1}\left(\widehat{\tau(A)}\right)$, the computational cost being $O(n^3)$. This cost is, of course, rather high, though no more so than Gaussian elimination. This would be a particularly interesting way to perturb our distance functional matrix, for the scalings embodied in Δ might reflect physical dimensions in the problem which it is desirable to perturb as little as possible. However, the computational expense, as well as the difficulty of extending the perturbed distance functional $\widehat{\Delta}$ so that it is defined for any points $x, y \in \mathbb{R}^d$, seems to limit this rather natural choice. [We might even interpolate $\sqrt{\widehat{\Delta}(x,y)}$ at the $n(n-1)/2$ points $\{(x_j, x_k) : j \neq k\}$ in $\mathbb{R}^d \times \mathbb{R}^d$, using a second radial basis function interpolant, but this would incur an $O(n^6)$ overhead.]

Fortunately, there is no particular need to go to the trouble of computing the nearest Euclidean distance matrix, in any sense, if the scalings chosen are only rough estimates. In this case, we can use the fact that, given any symmetric matrix M, the linear combination $M + \mu I$ is positive definite for all sufficiently large $\mu > 0$. In other words, we take $\tau^{-1}\left(\tau(A) + \mu I\right) = A + \mu\tau^{-1}(I)$. Further, it is not difficult to explicitly calculate the matrix $\tau^{-1}(I)$.

Lemma 2.6. *Let $M \in \text{Æ}_n$ be defined by $M_{jn} = 1/2$, for $1 \le j \le n-1$, and $M_{jk} = 1 - \delta_{jk}$, for $1 \le j, k \le n-1$ (δ_{jk} being the Kronecker delta). Then $\tau(M) = I$.*

Proof. By definition of the map τ,

$$\tau(M)_{jk} = M_{jn} + M_{kn} - M_{jk} = \frac{1}{2} + \frac{1}{2} - (1 - \delta_{jk}) = \delta_{jk}, \qquad 1 \le j, k \le n-1,$$

as required. □

Let us now summarise our findings formally.

Theorem 2.7. *Let $D \in \mathbb{R}^{n \times n}$ be the almost Euclidean matrix formed by the distance functional, that is,*

$$D_{jk} = \Delta(x_j, x_k), \qquad 1 \le j, k \le n.$$

(1) *The matrix $D(\mu) := D + \mu\tau^{-1}(I)$ is a Euclidean distance matrix for all sufficiently large $\mu > 0$.*
(2) *We can calculate the closest Euclidean distance matrix in the pulled-back Frobenius norm using the Matlab commands:*

```
[Q, Lambda] = eig(tau(D));
Dhat = tauinv(Q* max(Lambda, zeros(n))*Q');
```

However, it is not sufficient to merely calculate a perturbed distance functional matrix that is a Euclidean distance matrix. We must extend the definition of $\widehat{D}$ to obtain a new distance functional $\widehat{\Delta}$. This is particularly simple if we choose to form $D(\mu) = D + \mu\tau^{-1}(I)$. Specifically, we define $\widehat{\Delta} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ by

$$\widehat{\Delta}(x, x_n) = \Delta(x, x_n) + \mu/2, \qquad x \in \mathbb{R}^d \setminus \{x_n\},$$

and let

$$\widehat{\Delta}(x, y) = \Delta(x, y) + \mu$$

for any distinct points $x, y \in \mathbb{R}^d \setminus \{x_n\}$; of course, we define $\widehat{\Delta}(x, x) = 0$, for all $x \in \mathbb{R}^d$. As for choosing the constant μ, one simple way is to ensure that $D(\mu)$ is strictly diagonally dominant, that is

$$D(\mu)_{jj} > \sum_{k=1, k\neq j}^{n} |D(\mu)_{jk}|, \qquad \text{for } 1 \le j \le n-1.$$

In other words, we must have

$$2D_{jn} + \mu > \sum_{k=1, k\neq j}^{n} |D_{jn} + D_{kn} - D_{jk}|, \qquad 1 \le j \le n-1.$$

This justifies the choice specified in Theorem 1.5.

Finally, we observe that there are, in fact, infinitely many alternative linear bijections $\tau : Æ_n \to \mathrm{Symm}_{n-1}$ which satisfy $\tau(E_n) = P_{n-1}$, and this more general setting is considered by Gower [3], but τ has the virtue of simplicity. Nevertheless, such alternatives present an obvious topic of further research, together with the construction of continuous extensions of $\widehat{\Delta}$.

References

1. B. J. C. Baxter, *Conditionally positive functions and p-norm distance matrices*, Constructive Approximation **7** (1991), 427–440.
2. ———, *Positive definite functions on Hilbert space*, East Journal of Approximation **10** (2004), 269–274.
3. J. C. Gower, *Properties of Euclidean and non-Euclidean matrices*, Linear Algebra Appl **67** (1985), 81–97.
4. Nicholas J. Higham, *Matrix nearness problems and applications*, Applications of Matrix Theory (M. J. C. Gover and S. Barnett, eds.), Oxford University Press, 1989, pp. 1–27.
5. Y. C. Hon and E. J. Kansa, *Circumventing the ill-conditioning problem with multiquadric radial basis functions: applications to elliptic partial differential equations*, Comput. Math. Applic. **39** (2000), 123–127.
6. J. T. Kent K. V. Mardia and J. M. Bibby, *Multivariate analysis*, Academic Press, 1979.
7. M. D. Buhmann, *Radial Basis Functions: theory and implementations*, Cambridge University Press, 2003.
8. R. Mathar, *The best Euclidean fit to a given distance matrix in prescribed dimensions*, Linear Algebra Appl. **67** (1985), 1–6.
9. C. A. Micchelli, *Interpolation of scattered data: distance matrices and conditionally positive functions*, Constructive Approximation **2** (1986), 11–22.
10. F. J. Narcowich and J. D. Ward, *Norm estimates for the inverses of a general class of scattered-data radial function interpolation matrices*, J. Approx. Theory **69** (1992), 84–109.
11. I. J. Schoenberg, *Remarks to Maurice Fréchet's article 'sur la definition d'une classe d'espace distanciés vectoriellement applicable sur l'espace d'Hilbert.'*, Annals of Mathematics **36** (1935), 724–732.

12. ______, *On certain metric spaces arising from Euclidean space by a change of metric and their embedding in Hilbert space*, Annals of Mathematics **38** (1937), 787–793.
13. ______, *Metric spaces and completely monotone functions*, Annals of Mathematics **1938** (1938), 811–841.
14. ______, *Selected papers*, Birkhäuser, 1988.

School of Economics, Mathematics and Statistics, Birkbeck College, University of London, Malet Street, London WC1E 7HX, England
E-mail address: `b.baxter@bbk.ac.uk`