Schoenberg: from metric geometry to matrix positivity

Apoorva Khare
Indian Institute of Science

Eigenfunctions Seminar (with Gautam Bharali)
IISc, April 2019

Entrywise functions preserving positivity

Definitions:

1. A real symmetric matrix $A_{n \times n}$ is *positive semidefinite* if its quadratic form is so: $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$. (Hence $\sigma(A) \subset [0, \infty)$.)
2. Given $n \geq 1$ and $I \subset \mathbb{R}$, let $\mathbb{P}_n(I)$ denote the $n \times n$ positive (semidefinite) matrices, with entries in I. (Say $\mathbb{P}_n = \mathbb{P}_n(\mathbb{R})$.)

Entrywise functions preserving positivity

Definitions:

1. A real symmetric matrix $A_{n \times n}$ is *positive semidefinite* if its quadratic form is so: $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$. (Hence $\sigma(A) \subset [0, \infty)$.)
2. Given $n \geq 1$ and $I \subset \mathbb{R}$, let $\mathbb{P}_n(I)$ denote the $n \times n$ positive (semidefinite) matrices, with entries in I. (Say $\mathbb{P}_n = \mathbb{P}_n(\mathbb{R})$.)
3. A function $f : I \to \mathbb{R}$ acts *entrywise* on a matrix $A \in I^{n \times n}$ via: $f[A] := (f(a_{jk}))_{j,k=1}^n$.

Entrywise functions preserving positivity

Definitions:

1. A real symmetric matrix $A_{n\times n}$ is *positive semidefinite* if its quadratic form is so: $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$. (Hence $\sigma(A) \subset [0,\infty)$.)
2. Given $n \geq 1$ and $I \subset \mathbb{R}$, let $\mathbb{P}_n(I)$ denote the $n \times n$ positive (semidefinite) matrices, with entries in I. (Say $\mathbb{P}_n = \mathbb{P}_n(\mathbb{R})$.)
3. A function $f : I \to \mathbb{R}$ acts *entrywise* on a matrix $A \in I^{n\times n}$ via: $f[A] := (f(a_{jk}))_{j,k=1}^n$.

Problem: For which functions $f : I \to \mathbb{R}$ is it true that

$$f[A] \in \mathbb{P}_n \text{ for all } A \in \mathbb{P}_n(I)?$$

Entrywise functions preserving positivity

Definitions:

1. A real symmetric matrix $A_{n \times n}$ is *positive semidefinite* if its quadratic form is so: $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$. (Hence $\sigma(A) \subset [0, \infty)$.)
2. Given $n \geq 1$ and $I \subset \mathbb{R}$, let $\mathbb{P}_n(I)$ denote the $n \times n$ positive (semidefinite) matrices, with entries in I. (Say $\mathbb{P}_n = \mathbb{P}_n(\mathbb{R})$.)
3. A function $f : I \to \mathbb{R}$ acts *entrywise* on a matrix $A \in I^{n \times n}$ via: $f[A] := (f(a_{jk}))_{j,k=1}^n$.

Problem: For which functions $f : I \to \mathbb{R}$ is it true that

$$f[A] \in \mathbb{P}_n \text{ for all } A \in \mathbb{P}_n(I)?$$

- (Long history!) The Schur Product Theorem [Schur, *Crelle* 1911] says: If $A, B \in \mathbb{P}_n$, then so is $A \circ B := (a_{jk} b_{jk})$.

Entrywise functions preserving positivity

Definitions:

1. A real symmetric matrix $A_{n \times n}$ is *positive semidefinite* if its quadratic form is so: $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$. (Hence $\sigma(A) \subset [0, \infty)$.)
2. Given $n \geq 1$ and $I \subset \mathbb{R}$, let $\mathbb{P}_n(I)$ denote the $n \times n$ positive (semidefinite) matrices, with entries in I. (Say $\mathbb{P}_n = \mathbb{P}_n(\mathbb{R})$.)
3. A function $f : I \to \mathbb{R}$ acts *entrywise* on a matrix $A \in I^{n \times n}$ via: $f[A] := (f(a_{jk}))_{j,k=1}^n$.

Problem: For which functions $f : I \to \mathbb{R}$ is it true that

$$f[A] \in \mathbb{P}_n \text{ for all } A \in \mathbb{P}_n(I)?$$

- (Long history!) The Schur Product Theorem [Schur, *Crelle* 1911] says: If $A, B \in \mathbb{P}_n$, then so is $A \circ B := (a_{jk} b_{jk})$.
- As a consequence, $f(x) = x^k$ $(k \geq 0)$ preserves positivity on $\mathbb{P}_n$ for all n.

Entrywise functions preserving positivity

Definitions:

1. A real symmetric matrix $A_{n \times n}$ is *positive semidefinite* if its quadratic form is so: $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$. (Hence $\sigma(A) \subset [0, \infty)$.)
2. Given $n \geq 1$ and $I \subset \mathbb{R}$, let $\mathbb{P}_n(I)$ denote the $n \times n$ positive (semidefinite) matrices, with entries in I. (Say $\mathbb{P}_n = \mathbb{P}_n(\mathbb{R})$.)
3. A function $f : I \to \mathbb{R}$ acts *entrywise* on a matrix $A \in I^{n \times n}$ via: $f[A] := (f(a_{jk}))_{j,k=1}^n$.

Problem: For which functions $f : I \to \mathbb{R}$ is it true that

$$f[A] \in \mathbb{P}_n \text{ for all } A \in \mathbb{P}_n(I)?$$

- (Long history!) The Schur Product Theorem [Schur, *Crelle* 1911] says: If $A, B \in \mathbb{P}_n$, then so is $A \circ B := (a_{jk} b_{jk})$.
- As a consequence, $f(x) = x^k$ $(k \geq 0)$ preserves positivity on $\mathbb{P}_n$ for all n.
- (Pólya–Szegö, 1925): Taking sums and limits, if $f(x) = \sum_{k=0}^\infty c_k x^k$ is convergent and $c_k \geq 0$, then $f[-]$ preserves positivity.

Entrywise functions preserving positivity

Definitions:

1. A real symmetric matrix $A_{n \times n}$ is *positive semidefinite* if its quadratic form is so: $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$. (Hence $\sigma(A) \subset [0, \infty)$.)
2. Given $n \geq 1$ and $I \subset \mathbb{R}$, let $\mathbb{P}_n(I)$ denote the $n \times n$ positive (semidefinite) matrices, with entries in I. (Say $\mathbb{P}_n = \mathbb{P}_n(\mathbb{R})$.)
3. A function $f : I \to \mathbb{R}$ acts *entrywise* on a matrix $A \in I^{n \times n}$ via: $f[A] := (f(a_{jk}))_{j,k=1}^n$.

Problem: For which functions $f : I \to \mathbb{R}$ is it true that

$$f[A] \in \mathbb{P}_n \text{ for all } A \in \mathbb{P}_n(I)?$$

- (Long history!) The Schur Product Theorem [Schur, *Crelle* 1911] says: If $A, B \in \mathbb{P}_n$, then so is $A \circ B := (a_{jk} b_{jk})$.
- As a consequence, $f(x) = x^k$ $(k \geq 0)$ preserves positivity on $\mathbb{P}_n$ for all n.
- (Pólya–Szegö, 1925): Taking sums and limits, if $f(x) = \sum_{k=0}^\infty c_k x^k$ is convergent and $c_k \geq 0$, then $f[-]$ preserves positivity.

  Question: Anything else?

Schoenberg's theorem

Interestingly, the answer is **no**, if we want to preserve positivity in *all* dimensions:

Schoenberg's theorem

Interestingly, the answer is **no**, if we want to preserve positivity in *all* dimensions:

Theorem (Schoenberg, *Duke Math. J.* 1942; Rudin, *Duke Math. J.* 1959)

Suppose $I = (-1, 1)$ *and* $f : I \to \mathbb{R}$. *The following are equivalent:*

1. $f[A] \in \mathbb{P}_n$ *for all* $A \in \mathbb{P}_n(I)$ *and all* $n \geq 1$.
2. f *is analytic on* I *and has nonnegative Taylor coefficients.* *In other words,* $f(x) = \sum_{k=0}^{\infty} c_k x^k$ *on* $(-1, 1)$ *with all* $c_k \geq 0$.

Schoenberg's theorem

Interestingly, the answer is **no**, if we want to preserve positivity in *all* dimensions:

Theorem (Schoenberg, *Duke Math. J.* 1942; Rudin, *Duke Math. J.* 1959)

Suppose $I = (-1,1)$ *and* $f : I \to \mathbb{R}$. *The following are equivalent:*

1. $f[A] \in \mathbb{P}_n$ *for all* $A \in \mathbb{P}_n(I)$ *and all* $n \geq 1$.
2. f *is analytic on* I *and has nonnegative Taylor coefficients.* *In other words,* $f(x) = \sum_{k=0}^{\infty} c_k x^k$ *on* $(-1,1)$ *with all* $c_k \geq 0$.

- Schoenberg's result is the (harder) converse to that of his advisor: Schur.
- Vasudeva (1979) proved a variant, over $I = (0, \infty)$.

Schoenberg's theorem

Interestingly, the answer is **no**, if we want to preserve positivity in *all* dimensions:

Theorem (Schoenberg, *Duke Math. J.* 1942; Rudin, *Duke Math. J.* 1959)

Suppose $I = (-1, 1)$ *and* $f : I \to \mathbb{R}$*. The following are equivalent:*

1. $f[A] \in \mathbb{P}_n$ *for all* $A \in \mathbb{P}_n(I)$ *and all* $n \geq 1$.
2. f *is analytic on* I *and has nonnegative Taylor coefficients.*
   In other words, $f(x) = \sum_{k=0}^{\infty} c_k x^k$ *on* $(-1, 1)$ *with all* $c_k \geq 0$.

- Schoenberg's result is the (harder) converse to that of his advisor: Schur.
- Vasudeva (1979) proved a variant, over $I = (0, \infty)$.
- *Upshot:* Preserving positivity in all dimensions is a rigid condition $\rightsquigarrow$ implies real analyticity, absolute monotonicity…

Schoenberg's motivations: metric geometry

Endomorphisms of matrix spaces with positivity constraints related to:

- matrix monotone functions (Loewner)
- preservers of matrix properties (rank, inertia, ...)
- real-stable/hyperbolic polynomials (Borcea, Branden, Liggett, Marcus, Spielman, Srivastava...)
- positive definite functions (von Neumann, Bochner, Schoenberg ...)

Schoenberg's motivations: metric geometry

Endomorphisms of matrix spaces with positivity constraints related to:

- matrix monotone functions (Loewner)
- preservers of matrix properties (rank, inertia, ...)
- real-stable/hyperbolic polynomials (Borcea, Branden, Liggett, Marcus, Spielman, Srivastava...)
- positive definite functions (von Neumann, Bochner, Schoenberg ...)

Definition

$f : [0, \infty) \to \mathbb{R}$ is *positive definite* on a metric space (X, d) if
$[f(d(x_j, x_k))]_{j,k=1}^n \in \mathbb{P}_n, \quad$ for all $n \geq 1$ and all $x_1, \dots, x_n \in X$.

Schoenberg's motivations: metric geometry

Endomorphisms of matrix spaces with positivity constraints related to:

- matrix monotone functions (Loewner)
- preservers of matrix properties (rank, inertia, ...)
- real-stable/hyperbolic polynomials (Borcea, Branden, Liggett, Marcus, Spielman, Srivastava...)
- positive definite functions (von Neumann, Bochner, Schoenberg ...)

Definition

$f : [0, \infty) \to \mathbb{R}$ is *positive definite* on a metric space (X, d) if
$[f(d(x_j, x_k))]_{j,k=1}^n \in \mathbb{P}_n, \quad$ for all $n \geq 1$ and all $x_1, \dots, x_n \in X$.

Plan for rest of the talk: Discuss the path from metric geometry, through positive definite functions, to Schoenberg's theorem.

Distance geometry

How did the study of positivity and its preservers begin?

Distance geometry

How did the study of positivity and its preservers begin?

In the 1900s, the notion of a *metric space* emerged from the works of Fréchet and Hausdorff. . .

- Now ubiquitous in science (mathematics, physics, economics, statistics, computer science. . .).

How did the study of positivity and its preservers begin?

In the 1900s, the notion of a *metric space* emerged from the works of Fréchet and Hausdorff...

- Now ubiquitous in science (mathematics, physics, economics, statistics, computer science...).
- Fréchet [*Math. Ann.* 1910]. If (X, d) is a metric space with $|X| = n + 1$, then (X, d) isometrically embeds into $(\mathbb{R}^n, \ell_\infty)$.

How did the study of positivity and its preservers begin?

In the 1900s, the notion of a *metric space* emerged from the works of Fréchet and Hausdorff...

- Now ubiquitous in science (mathematics, physics, economics, statistics, computer science...).
- Fréchet [*Math. Ann.* 1910]. If (X, d) is a metric space with $|X| = n + 1$, then (X, d) isometrically embeds into $(\mathbb{R}^n, \ell_\infty)$.
- This avenue of work led to the exploration of metric space embeddings. Natural question: *Which metric spaces isometrically embed into* ***Euclidean space****?*

Euclidean metric spaces and positive matrices

Which metric spaces isometrically embed into a Euclidean space?

Euclidean metric spaces and positive matrices

Which metric spaces isometrically embed into a Euclidean space?

- Menger [*Amer. J. Math.* 1931] and Fréchet [*Ann. of Math.* 1935] provided characterizations.

Euclidean metric spaces and positive matrices

Which metric spaces isometrically embed into a Euclidean space?

- Menger [*Amer. J. Math.* 1931] and Fréchet [*Ann. of Math.* 1935] provided characterizations.
- Reformulated by Schoenberg, using... matrix positivity!

Theorem (Schoenberg, *Ann. of Math.* 1935)

Fix integers $n, r \geq 1$, and a finite set $X = \{x_0, \dots, x_n\}$ together with a metric d on X. Then (X, d) isometrically embeds into $\mathbb{R}^r$ (with the Euclidean distance/norm) but not into $\mathbb{R}^{r-1}$ if and only if the $n \times n$ matrix

$$A := (d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2)_{j,k=1}^n$$

is positive semidefinite of rank r.

Connects metric geometry and matrix positivity.

Schoenberg: from metric geometry to matrix positivity

Sketch of one implication: If (X, d) isometrically embeds into $(\mathbb{R}^r, \| \cdot \|)$, then

$$\begin{aligned} & d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2 \\ = & \|x_0 - x_j\|^2 + \|x_0 - x_k\|^2 - \|(x_0 - x_j) - (x_0 - x_k)\|^2 \\ = & 2\langle x_0 - x_j, x_0 - x_k \rangle. \end{aligned}$$

But then the matrix A above, is the Gram matrix of a set of vectors in $\mathbb{R}^r$, hence is positive semidefinite, of rank $\leq r$.

Schoenberg: from metric geometry to matrix positivity

Sketch of one implication: If (X, d) isometrically embeds into $(\mathbb{R}^r, \|\cdot\|)$, then

$$\begin{aligned} & d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2 \\ = & \|x_0 - x_j\|^2 + \|x_0 - x_k\|^2 - \|(x_0 - x_j) - (x_0 - x_k)\|^2 \\ = & 2\langle x_0 - x_j, x_0 - x_k \rangle. \end{aligned}$$

But then the matrix A above, is the Gram matrix of a set of vectors in $\mathbb{R}^r$, hence is positive semidefinite, of rank $\leq r$. In fact the rank is exactly r. □

Sketch of one implication: If (X, d) isometrically embeds into $(\mathbb{R}^r, \|\cdot\|)$, then

$$\begin{aligned} & d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2 \\ = & \|x_0 - x_j\|^2 + \|x_0 - x_k\|^2 - \|(x_0 - x_j) - (x_0 - x_k)\|^2 \\ = & 2\langle x_0 - x_j, x_0 - x_k \rangle. \end{aligned}$$

But then the matrix A above, is the Gram matrix of a set of vectors in $\mathbb{R}^r$, hence is positive semidefinite, of rank $\leq r$. In fact the rank is exactly r. □

- Also observe: the matrix $A := (d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2)_{j,k=1}^n$ is positive semidefinite,
  if and only if the matrix $A'_{(n+1)\times(n+1)} := (-d(x_j, x_k)^2)_{j,k=0}^n$ is *conditionally positive semidefinite*: $u^T A' u \geq 0$ whenever $\sum_{j=0}^n u_j = 0$.

Sketch of one implication: If (X, d) isometrically embeds into $(\mathbb{R}^r, \|\cdot\|)$, then

$$\begin{aligned} & d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2 \\ = & \|x_0 - x_j\|^2 + \|x_0 - x_k\|^2 - \|(x_0 - x_j) - (x_0 - x_k)\|^2 \\ = & 2\langle x_0 - x_j, x_0 - x_k \rangle. \end{aligned}$$

But then the matrix A above, is the Gram matrix of a set of vectors in $\mathbb{R}^r$, hence is positive semidefinite, of rank $\leq r$. In fact the rank is exactly r. □

- Also observe: the matrix $A := (d(x_0, x_j)^2 + d(x_0, x_k)^2 - d(x_j, x_k)^2)_{j,k=1}^n$ is positive semidefinite,
  if and only if the matrix $A'_{(n+1)\times(n+1)} := (-d(x_j, x_k)^2)_{j,k=0}^n$ is *conditionally positive semidefinite*: $u^T A' u \geq 0$ whenever $\sum_{j=0}^n u_j = 0$.
- This is how positive / conditionally positive matrices emerged from metric geometry.

Distance transforms: positive definite functions

As we saw, applying the function $-x^2$ *entrywise* sends any distance matrix from Euclidean space, to a conditionally positive semidefinite matrix A'.

What operations send distance matrices to positive semidefinite matrices?

Distance transforms: positive definite functions

As we saw, applying the function $-x^2$ *entrywise* sends any distance matrix from Euclidean space, to a conditionally positive semidefinite matrix A'.

What operations send distance matrices to positive semidefinite matrices? These are the *positive definite functions*. **Example:** Gaussian kernel:

Theorem (Schoenberg, *Trans. AMS* 1938)

The function $f(x) = \exp(-x^2)$ is positive definite on $\mathbb{R}^r$, for all $r \geq 1$.

Schoenberg showed this using Bochner's theorem on $\mathbb{R}^r$, and the fact that the Gaussian function is its own Fourier transform (up to constants).

Distance transforms: positive definite functions

As we saw, applying the function $-x^2$ *entrywise* sends any distance matrix from Euclidean space, to a conditionally positive semidefinite matrix A'.

What operations send distance matrices to positive semidefinite matrices? These are the *positive definite functions*. **Example:** Gaussian kernel:

Theorem (Schoenberg, *Trans. AMS* 1938)

The function $f(x) = \exp(-x^2)$ is positive definite on $\mathbb{R}^r$, for all $r \geq 1$.

Schoenberg showed this using Bochner's theorem on $\mathbb{R}^r$, and the fact that the Gaussian function is its own Fourier transform (up to constants).

Alternate proof (K.):
(1) An observation of Gantmakher and Krein(?): *Generalized Vandermonde matrices are totally positive.* In other words, if $0 < y_1 < \cdots < y_n$ and $x_1 < \cdots < x_n$ in $\mathbb{R}$, then $\det(y_j^{x_k})_{j,k=1}^n$ is positive.

(2) A result by Pólya: *The Gaussian kernel is positive definite on $\mathbb{R}^1$.* Indeed,

$$\left(\exp(-(x_j - x_k)^2)\right)_{j,k=1}^n = \operatorname{diag}(e^{-x_j^2}) \times \left(\exp(2x_j x_k)\right)_{j,k=1}^n \times \operatorname{diag}(e^{-x_k^2}).$$

(3) A result of Schur: *The Schur product theorem* implies the result for $\mathbb{R}^r$. □

Metric embeddings via the Gaussian kernel

This implies the 'only if' part of the following result:

Theorem (Schoenberg, *Trans. AMS* 1938)

A finite metric space (X, d) *with* $X = \{x_0, \dots, x_n\}$ *embeds isometrically into* $\mathbb{R}^r$ *for some* $r > 0$ *(which turns out to be at most* n*), if and only if for all* $\lambda > 0$, *the* $(n+1) \times (n+1)$ *matrix* X_λ *with* (j, k) *entry*

$$(X_\lambda)_{j,k} := \exp(-\lambda^2 d(x_j, x_k)^2), \qquad 0 \leq j, k \leq n$$

is positive semidefinite. (I.e., $\exp(-\lambda^2 x^2)$ *is positive definite on* X*.)*

Note again the connection between metric geometry and matrix positivity.

Metric embeddings via the Gaussian kernel

This implies the 'only if' part of the following result:

Theorem (Schoenberg, *Trans. AMS* 1938)

A finite metric space (X, d) *with* $X = \{x_0, \dots, x_n\}$ *embeds isometrically into* $\mathbb{R}^r$ *for some* $r > 0$ *(which turns out to be at most* n*), if and only if for all* $\lambda > 0$*, the* $(n+1) \times (n+1)$ *matrix* X_λ *with* (j, k) *entry*

$$(X_\lambda)_{j,k} := \exp(-\lambda^2 d(x_j, x_k)^2), \qquad 0 \leq j, k \leq n$$

is positive semidefinite. (I.e., $\exp(-\lambda^2 x^2)$ *is positive definite on* X*.)*

Note again the connection between metric geometry and matrix positivity.

Proof of 'if' part:

We only need that X_λ is conditionally positive. If $\sum_{j \geq 0} u_j = 0$, then expanding $u^T X_\lambda u \geq 0$ as a power series in $\lambda^2 \ll 1$, the first two leading terms are:

$$\lambda^0 \sum_{j,k=0}^{n} u_j u_k = \Big(\sum_{j \geq 0} u_j\Big)^2 = 0, \qquad -\lambda^2 \sum_{j,k=0}^{n} u_j u_k d(x_j, x_k)^2.$$

Metric embeddings via the Gaussian kernel

This implies the 'only if' part of the following result:

Theorem (Schoenberg, *Trans. AMS* 1938)

A finite metric space (X, d) *with* $X = \{x_0, \dots, x_n\}$ *embeds isometrically into* $\mathbb{R}^r$ *for some* $r > 0$ *(which turns out to be at most* n*), if and only if for all* $\lambda > 0$, *the* $(n+1) \times (n+1)$ *matrix* X_λ *with* (j, k) *entry*

$$(X_\lambda)_{j,k} := \exp(-\lambda^2 d(x_j, x_k)^2), \qquad 0 \le j, k \le n$$

is positive semidefinite. (I.e., $\exp(-\lambda^2 x^2)$ *is positive definite on* X*.)*

Note again the connection between metric geometry and matrix positivity.

Proof of 'if' part:
We only need that X_λ is conditionally positive. If $\sum_{j \ge 0} u_j = 0$, then expanding $u^T X_\lambda u \ge 0$ as a power series in $\lambda^2 \ll 1$, the first two leading terms are:

$$\lambda^0 \sum_{j,k=0}^{n} u_j u_k = \Big(\sum_{j \ge 0} u_j\Big)^2 = 0, \qquad -\lambda^2 \sum_{j,k=0}^{n} u_j u_k d(x_j, x_k)^2.$$

Thus the leading coefficient (of λ^2) is non-negative, so $A' = (-d(x_j, x_k)^2)_{j,k=0}^n$ is conditionally positive.

This implies the 'only if' part of the following result:

Theorem (Schoenberg, *Trans. AMS* 1938)

A finite metric space (X, d) with $X = \{x_0, \dots, x_n\}$ embeds isometrically into $\mathbb{R}^r$ for some $r > 0$ (which turns out to be at most n), if and only if for all $\lambda > 0$, the $(n+1) \times (n+1)$ matrix X_λ with (j, k) entry

$$(X_\lambda)_{j,k} := \exp(-\lambda^2 d(x_j, x_k)^2), \qquad 0 \leq j, k \leq n$$

is positive semidefinite. (I.e., $\exp(-\lambda^2 x^2)$ is positive definite on X.)

Note again the connection between metric geometry and matrix positivity.

Proof of 'if' part:
We only need that X_λ is conditionally positive. If $\sum_{j \geq 0} u_j = 0$, then expanding $u^T X_\lambda u \geq 0$ as a power series in $\lambda^2 \ll 1$, the first two leading terms are:

$$\lambda^0 \sum_{j,k=0}^{n} u_j u_k = \Big(\sum_{j \geq 0} u_j\Big)^2 = 0, \qquad -\lambda^2 \sum_{j,k=0}^{n} u_j u_k d(x_j, x_k)^2.$$

Thus the leading coefficient (of λ^2) is non-negative, so $A' = (-d(x_j, x_k)^2)_{j,k=0}^n$ is conditionally positive. Now apply Schoenberg's 1935 result. □

Spherical embeddings, via positive definite maps

The previous result says: Euclidean spaces $\mathbb{R}^r$, or their direct limit $\mathbb{R}^\infty$ (called *Hilbert space* by Schoenberg) are characterized by the property that the maps

$$\exp(-\lambda^2 x^2), \qquad \lambda \in (0, \epsilon)$$

are all positive definite on each (finite) metric subspace.
(As we saw, such a characterization holds for each $\epsilon > 0$.)

Spherical embeddings, via positive definite maps

The previous result says: Euclidean spaces $\mathbb{R}^r$, or their direct limit $\mathbb{R}^\infty$ (called *Hilbert space* by Schoenberg) are characterized by the property that the maps

$$\exp(-\lambda^2 x^2), \qquad \lambda \in (0, \epsilon)$$

are all positive definite on each (finite) metric subspace.
(As we saw, such a characterization holds for each $\epsilon > 0$.)

What about distinguished subsets of $\mathbb{R}^r$ or of $\mathbb{R}^\infty$? Can one find similar families of functions for them?

Spherical embeddings, via positive definite maps

The previous result says: Euclidean spaces $\mathbb{R}^r$, or their direct limit $\mathbb{R}^\infty$ (called *Hilbert space* by Schoenberg) are characterized by the property that the maps

$$\exp(-\lambda^2 x^2), \qquad \lambda \in (0, \epsilon)$$

are all positive definite on each (finite) metric subspace.
(As we saw, such a characterization holds for each $\epsilon > 0$.)

What about distinguished subsets of $\mathbb{R}^r$ or of $\mathbb{R}^\infty$? Can one find similar families of functions for them?

Schoenberg explored this question for **spheres**: $S^{r-1} \subset \mathbb{R}^r$ and $S^\infty \subset \mathbb{R}^\infty$.
It turns out, the characterization now involves a *single* function!

The previous result says: Euclidean spaces $\mathbb{R}^r$, or their direct limit $\mathbb{R}^\infty$ (called *Hilbert space* by Schoenberg) are characterized by the property that the maps

$$\exp(-\lambda^2 x^2), \qquad \lambda \in (0, \epsilon)$$

are all positive definite on each (finite) metric subspace.
(As we saw, such a characterization holds for each $\epsilon > 0$.)

What about distinguished subsets of $\mathbb{R}^r$ or of $\mathbb{R}^\infty$? Can one find similar families of functions for them?

Schoenberg explored this question for **spheres**: $S^{r-1} \subset \mathbb{R}^r$ and $S^\infty \subset \mathbb{R}^\infty$.
It turns out, the characterization now involves a *single* function!

This is the *cosine function*.

Spherical embeddings via cosines

Notice that the Hilbert sphere S^∞ (hence every subspace such as S^{r-1}) has a rotation-invariant distance – *arc-length* along a great circle:

$$d(x,y) := \sphericalangle(x,y) = \arccos\langle x,y\rangle, \qquad x,y \in S^\infty.$$

Hence applying $\cos[-]$ entrywise to any distance matrix on S^∞ yields:

$$\cos[(d(x_j,x_k))_{j,k\geq 0}] = (\langle x_j,x_k\rangle)_{j,k\geq 0},$$

and this is a Gram matrix, so positive semidefinite.

Spherical embeddings via cosines

Notice that the Hilbert sphere S^∞ (hence every subspace such as S^{r-1}) has a rotation-invariant distance – *arc-length* along a great circle:

$$d(x,y) := \sphericalangle(x,y) = \arccos\langle x,y\rangle, \qquad x,y \in S^\infty.$$

Hence applying $\cos[-]$ entrywise to any distance matrix on S^∞ yields:

$$\cos[(d(x_j,x_k))_{j,k\geq 0}] = (\langle x_j,x_k\rangle)_{j,k\geq 0},$$

and this is a Gram matrix, so positive semidefinite. Moreover, if $X \hookrightarrow S^\infty$ then X must have diameter at most $\operatorname{diam} S^\infty = \pi$. This shows one half of:

Theorem (Schoenberg, *Ann. of Math.* 1935)

A finite metric space (X,d) embeds isometrically into the Hilbert sphere S^∞ if and only if (a) $\cos(x)$ is positive definite on X, and (b) $\operatorname{diam} X \leq \pi$.

Spherical embeddings via cosines

Notice that the Hilbert sphere S^∞ (hence every subspace such as S^{r-1}) has a rotation-invariant distance – *arc-length* along a great circle:

$$d(x,y) := \sphericalangle(x,y) = \arccos\langle x,y\rangle, \qquad x,y \in S^\infty.$$

Hence applying $\cos[-]$ entrywise to any distance matrix on S^∞ yields:

$$\cos[(d(x_j,x_k))_{j,k\geq 0}] = (\langle x_j,x_k\rangle)_{j,k\geq 0},$$

and this is a Gram matrix, so positive semidefinite. Moreover, if $X \hookrightarrow S^\infty$ then X must have diameter at most $\operatorname{diam} S^\infty = \pi$. This shows one half of:

Theorem (Schoenberg, *Ann. of Math.* 1935)

A finite metric space (X,d) embeds isometrically into the Hilbert sphere S^∞ if and only if (a) $\cos(x)$ is positive definite on X, and (b) $\operatorname{diam} X \leq \pi$.

Proof of 'if' part: If $A := (\cos d(x_j,x_k))_{j,k=0}^n$ is positive semidefinite, write $A = B^T B$ for some $B_{r\times(n+1)}$ of rank $r = \operatorname{rank}(A)$.

- Let $y_0,\dots,y_n$ denote the columns of B. Then $y_j \in S^{r-1} \subset S^\infty$.
- Now check that $x_j \mapsto y_j$ is an isometric embedding : $X \hookrightarrow S^{r-1}$. $\square$

Positive definite functions on spheres

These results characterize $\mathbb{R}^\infty$ and S^∞ in terms of positive definite functions.

At the same time (1930s), Bochner proved his famous theorem classifying all positive definite functions on Euclidean space [*Math. Ann.* 1933].
Simultaneously generalized in 1940 by Weil, Povzner, and Raikov to arbitrary locally compact abelian groups.

Positive definite functions on spheres

These results characterize $\mathbb{R}^\infty$ and S^∞ in terms of positive definite functions.

At the same time (1930s), Bochner proved his famous theorem classifying all positive definite functions on Euclidean space [*Math. Ann.* 1933].
Simultaneously generalized in 1940 by Weil, Povzner, and Raikov to arbitrary locally compact abelian groups.

After understanding that $\cos(\cdot)$ is positive definite on S^∞, Schoenberg was interested in classifying *positive definite functions on spheres*.
This is the main result – and the title! – of his 1942 paper:

Positive definite functions on spheres

These results characterize $\mathbb{R}^\infty$ and S^∞ in terms of positive definite functions.

At the same time (1930s), Bochner proved his famous theorem classifying all positive definite functions on Euclidean space [*Math. Ann.* 1933].
Simultaneously generalized in 1940 by Weil, Povzner, and Raikov to arbitrary locally compact abelian groups.

After understanding that $\cos(\cdot)$ is positive definite on S^∞, Schoenberg was interested in classifying *positive definite functions on spheres*.
This is the main result – and the title! – of his 1942 paper:

Theorem (Schoenberg, *Duke Math. J.* 1942)

Suppose $f : [-1,1] \to \mathbb{R}$ *is continuous, and* $r \geq 2$. *Then* $f(\cos \cdot)$ *is positive definite on the unit sphere* $S^{r-1} \subset \mathbb{R}^r$ *if and only if*

$$f(\cdot) = \sum_{k \geq 0} a_k C_k^{(\frac{r-2}{2})}(\cdot) \qquad \text{for some } a_k \geq 0,$$

where $C_k^{(\lambda)}(\cdot)$ *are the ultraspherical / Gegenbauer / Chebyshev polynomials.*

Positive definite functions on spheres

These results characterize $\mathbb{R}^\infty$ and S^∞ in terms of positive definite functions.

At the same time (1930s), Bochner proved his famous theorem classifying all positive definite functions on Euclidean space [*Math. Ann.* 1933]. Simultaneously generalized in 1940 by Weil, Povzner, and Raikov to arbitrary locally compact abelian groups.

After understanding that $\cos(\cdot)$ is positive definite on S^∞, Schoenberg was interested in classifying *positive definite functions on spheres*.
This is the main result – and the title! – of his 1942 paper:

Theorem (Schoenberg, *Duke Math. J.* 1942)

Suppose $f : [-1,1] \to \mathbb{R}$ *is continuous, and* $r \geq 2$. *Then* $f(\cos \cdot)$ *is positive definite on the unit sphere* $S^{r-1} \subset \mathbb{R}^r$ *if and only if*

$$f(\cdot) = \sum_{k \geq 0} a_k C_k^{(\frac{r-2}{2})}(\cdot) \qquad \text{for some } a_k \geq 0,$$

where $C_k^{(\lambda)}(\cdot)$ *are the ultraspherical / Gegenbauer / Chebyshev polynomials.*

Also follows from Bochner's work on compact homogeneous spaces [*Ann. of Math.* 1941] – but Schoenberg proved it directly with less 'heavy' machinery.

From spheres to correlation matrices

- Any Gram matrix of vectors $x_j \in S^{r-1}$ is the same as a rank $\leq r$ correlation matrix $A = (a_{jk})_{j,k=1}^n$, i.e.,

$$A = \begin{pmatrix} 1 & & * & \\ & 1 & & \\ * & & 1 & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} — & x_1^T & — \\ — & x_2^T & — \\ & \vdots & \\ — & x_n^T & — \end{pmatrix} \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_n \\ | & | & & | \end{pmatrix} = (\langle x_j, x_k \rangle)_{j,k=1}^n.$$

From spheres to correlation matrices

- Any Gram matrix of vectors $x_j \in S^{r-1}$ is the same as a rank $\leq r$ correlation matrix $A = (a_{jk})_{j,k=1}^n$, i.e.,

$$A = \begin{pmatrix} 1 & & * & \\ & 1 & & \\ * & & 1 & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} — & x_1^T & — \\ — & x_2^T & — \\ & \vdots & \\ — & x_n^T & — \end{pmatrix} \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_n \\ | & | & & | \end{pmatrix} = (\langle x_j, x_k \rangle)_{j,k=1}^n.$$

- So,

$$\begin{aligned} f(\cos \cdot) \text{ positive definite on } S^{r-1} \quad &\iff \quad (f(\cos d(x_j, x_k)))_{j,k=1}^n \in \mathbb{P}_n \\ &\iff \quad (f(\langle x_j, x_k \rangle))_{j,k=1}^n \in \mathbb{P}_n \\ &\iff \quad (f(a_{jk}))_{j,k=1}^n \in \mathbb{P}_n \ \forall n \geq 1, \end{aligned}$$

From spheres to correlation matrices

- Any Gram matrix of vectors $x_j \in S^{r-1}$ is the same as a rank $\leq r$ correlation matrix $A = (a_{jk})_{j,k=1}^n$, i.e.,

$$A = \begin{pmatrix} 1 & & * & \\ & 1 & & \\ & * & 1 & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} - & x_1^T & - \\ - & x_2^T & - \\ & \vdots & \\ - & x_n^T & - \end{pmatrix} \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_n \\ | & | & & | \end{pmatrix} = (\langle x_j, x_k \rangle)_{j,k=1}^n.$$

- So,

$$\begin{aligned} f(\cos \cdot) \text{ positive definite on } S^{r-1} \quad & \iff \quad (f(\cos d(x_j, x_k)))_{j,k=1}^n \in \mathbb{P}_n \\ & \iff \quad (f(\langle x_j, x_k \rangle))_{j,k=1}^n \in \mathbb{P}_n \\ & \iff \quad (f(a_{jk}))_{j,k=1}^n \in \mathbb{P}_n \ \forall n \geq 1, \end{aligned}$$

i.e., f preserves positivity on correlation matrices of rank $\leq r$.

From spheres to correlation matrices

- Any Gram matrix of vectors $x_j \in S^{r-1}$ is the same as a rank $\leq r$ correlation matrix $A = (a_{jk})_{j,k=1}^n$, i.e.,

$$A = \begin{pmatrix} 1 & & * & \\ & 1 & & \\ & & 1 & \\ * & & & 1 \end{pmatrix} = \begin{pmatrix} - & x_1^T & - \\ - & x_2^T & - \\ & \vdots & \\ - & x_n^T & - \end{pmatrix} \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_n \\ | & | & & | \end{pmatrix} = (\langle x_j, x_k \rangle)_{j,k=1}^n.$$

- So,

$$\begin{aligned} f(\cos \cdot) \text{ positive definite on } S^{r-1} \quad & \iff \quad (f(\cos d(x_j, x_k)))_{j,k=1}^n \in \mathbb{P}_n \\ & \iff \quad (f(\langle x_j, x_k \rangle))_{j,k=1}^n \in \mathbb{P}_n \\ & \iff \quad (f(a_{jk}))_{j,k=1}^n \in \mathbb{P}_n \ \forall n \geq 1, \end{aligned}$$

  i.e., f preserves positivity on correlation matrices of rank $\leq r$.

- If instead $r = \infty$, such a result would classify the entrywise positivity preservers on all correlation matrices.

From spheres to correlation matrices

- Any Gram matrix of vectors $x_j \in S^{r-1}$ is the same as a rank $\leq r$ correlation matrix $A = (a_{jk})_{j,k=1}^n$, i.e.,

$$A = \begin{pmatrix} 1 & & * & \\ & 1 & & \\ * & & 1 & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} — & x_1^T & — \\ — & x_2^T & — \\ & \vdots & \\ — & x_n^T & — \end{pmatrix} \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_n \\ | & | & & | \end{pmatrix} = (\langle x_j, x_k \rangle)_{j,k=1}^n.$$

- So,

$$\begin{aligned} f(\cos \cdot) \text{ positive definite on } S^{r-1} \quad & \iff \quad (f(\cos d(x_j, x_k)))_{j,k=1}^n \in \mathbb{P}_n \\ & \iff \quad (f(\langle x_j, x_k \rangle))_{j,k=1}^n \in \mathbb{P}_n \\ & \iff \quad (f(a_{jk}))_{j,k=1}^n \in \mathbb{P}_n \ \forall n \geq 1, \end{aligned}$$

  i.e., f preserves positivity on correlation matrices of rank $\leq r$.

- If instead $r = \infty$, such a result would classify the entrywise positivity preservers on all correlation matrices. Interestingly, 70 years later the subject has acquired renewed interest because of its immediate impact in high-dimensional covariance estimation, in several applied fields.

Schoenberg's theorem on positivity preservers

And indeed, Schoenberg did make the leap from S^{r-1} to S^∞:

Theorem (Schoenberg, *Duke Math. J.* 1942)

Suppose $f : [-1,1] \to \mathbb{R}$ *is continuous. Then* $f(\cos \cdot)$ *is positive definite on the Hilbert sphere* $S^\infty \subset \mathbb{R}^\infty$ *if and only if*

$$f(\cos\theta) = \sum_{k \geq 0} c_k \cos^k \theta,$$

where $c_k \geq 0 \ \forall k$ *are such that* $\sum_k c_k < \infty$.

Schoenberg's theorem on positivity preservers

And indeed, Schoenberg did make the leap from S^{r-1} to S^∞:

Theorem (Schoenberg, *Duke Math. J.* 1942)

Suppose $f : [-1,1] \to \mathbb{R}$ is continuous. Then $f(\cos \cdot)$ is positive definite on the Hilbert sphere $S^\infty \subset \mathbb{R}^\infty$ if and only if

$$f(\cos\theta) = \sum_{k \geq 0} c_k \cos^k \theta,$$

where $c_k \geq 0 \ \forall k$ are such that $\sum_k c_k < \infty$.

Notice that $\cos^k \theta$ is positive definite on S^∞ for each $k \geq 0$, by the Schur product theorem.

Freeing this result from the sphere context, one obtains Schoenberg's theorem on entrywise positivity preservers.

Schoenberg's theorem on positivity preservers

And indeed, Schoenberg did make the leap from S^{r-1} to S^∞:

Theorem (Schoenberg, *Duke Math. J.* 1942)

Suppose $f : [-1,1] \to \mathbb{R}$ *is continuous. Then* $f(\cos \cdot)$ *is positive definite on the Hilbert sphere* $S^\infty \subset \mathbb{R}^\infty$ *if and only if*

$$f(\cos\theta) = \sum_{k \geq 0} c_k \cos^k \theta,$$

where $c_k \geq 0 \ \forall k$ *are such that* $\sum_k c_k < \infty$.

Notice that $\cos^k \theta$ is positive definite on S^∞ for each $k \geq 0$, by the Schur product theorem.

Freeing this result from the sphere context, one obtains Schoenberg's theorem on entrywise positivity preservers.

For more information: *A panorama of positivity* – available on arXiv. (Dec. 2018 survey by A. Belton, D. Guillot, A.K., and M. Putinar.)