MATH 590: Meshfree Methods
Chapter 2 — Part 2: Integral Characterizations of Positive Definite Functions

Greg Fasshauer

Department of Applied Mathematics
Illinois Institute of Technology

Fall 2014

Outline

1. Integral Characterizations for (Strictly) Positive Definite Functions

2. Positive Definite Radial Functions

We summarize some facts about integral characterizations of positive definite functions.

They were established in the 1930s by Bochner and Schoenberg.

We will also mention more recent extensions to strictly positive definite functions that are needed for the scattered data interpolation problem.

Many more details in [Wen05].

We start with a very brief discussion of concepts from measure theory and integral transforms that appear in the results later on.

Bochner's theorem below is formulated in terms of *Borel measures*.

Definition

Let X be an arbitrary set and denote by $\mathcal{P}(x)$ the set of all subsets of X. A subset $\mathcal{A}$ of $\mathcal{P}(X)$ is called a σ-algebra on X if

(1) $X \in \mathcal{A}$,

(2) $A \in \mathcal{A}$ implies that its complement (in X) is also contained in $\mathcal{A}$,

(3) $A_i \in \mathcal{A}$, $i \in \mathbb{N}$, implies that the union of these sets belongs to $\mathcal{A}$.

Definition

Given an arbitrary set X and a σ-algebra $\mathcal{A}$ of subsets of X, a measure on $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that

(1) $\mu(\emptyset) = 0$,

(2) for any sequence $\{A_i\}$ of disjoint sets in $\mathcal{A}$ we have

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

Definition

If X is a topological space, and $\mathcal{O}$ is the collection of open sets in X, then the σ-algebra generated by $\mathcal{O}$ is called the Borel σ-algebra and denoted by $\mathcal{B}(X)$.

If in addition X is a Hausdorff space[a], then a measure μ defined on $\mathcal{B}(X)$ that satisfies $\mu(K) < \infty$ for all compact sets $K \subseteq X$ is called a Borel measure.

The carrier of a Borel measure is given by the set

$$X \setminus \{O : \ O \in \mathcal{O} \text{ and } \mu(O) = 0\}.$$

---

[a]X is a Hausdorff space if any two distinct points of X can be separated by open sets

Definition

The Fourier transform of $f \in L_1(\mathbb{R}^d)$ is given by

$$\hat{f}(\boldsymbol{\omega}) = \frac{1}{\sqrt{(2\pi)^d}} \int_{\mathbb{R}^d} f(\boldsymbol{t}) e^{-i\boldsymbol{\omega} \cdot \boldsymbol{t}} d\boldsymbol{t}, \qquad \boldsymbol{\omega} \in \mathbb{R}^d,$$

and its inverse Fourier transform is given by

$$\check{f}(\boldsymbol{t}) = \frac{1}{\sqrt{(2\pi)^d}} \int_{\mathbb{R}^d} f(\boldsymbol{\omega}) e^{i\boldsymbol{t} \cdot \boldsymbol{\omega}} d\boldsymbol{\omega}, \qquad \boldsymbol{t} \in \mathbb{R}^d.$$

This definition from [Rud91] is based on angular frequency and is used in, e.g., [Wen05, SS02].

Another, just as common, definition from [SW71] uses ordinary frequency, i.e.,

$$\hat{f}_{SW}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\boldsymbol{t}) e^{-2\pi i \boldsymbol{\omega} \cdot \boldsymbol{t}} d\boldsymbol{t} = \sqrt{(2\pi)^d} \hat{f}(2\pi\boldsymbol{\omega}).$$

It appears, e.g., in [Buh03, CL99].

Similarly, we can define the Fourier transform of a finite (signed) measure μ on $\mathbb{R}^d$ by

$$\hat{\mu}(\boldsymbol{\omega}) = \frac{1}{\sqrt{(2\pi)^d}} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}\cdot\boldsymbol{t}} d\mu(\boldsymbol{t}), \qquad \boldsymbol{\omega} \in \mathbb{R}^d.$$

Since we are frequently interested in positive definite radial functions, we note that the Fourier transform of a radial function is again radial:

Theorem

Let $\Phi \in L_1(\mathbb{R}^d)$ be continuous and radial, i.e., $\Phi(\boldsymbol{x}) = \varphi(\|\boldsymbol{x}\|)$. Then its Fourier transform $\hat{\Phi}$ is also radial, i.e., $\hat{\Phi}(\boldsymbol{\omega}) = \mathcal{F}_d\varphi(\|\boldsymbol{\omega}\|)$ with

$$\mathcal{F}_d\varphi(r) = \frac{1}{\sqrt{r^{d-2}}} \int_0^\infty \varphi(t) t^{\frac{d}{2}} J_{\frac{d-2}{2}}(rt) \mathrm{d}t,$$

where $J_{\frac{d-2}{2}}$ is the classical Bessel function of the first kind of order $\frac{d-2}{2}$.

Proof.

See [Wen05].   □

This transform is also called Fourier-Bessel transform or Hankel transform.

Remark

The Hankel inversion theorem [Sne72] ensures that the Fourier transform for radial functions is its own inverse, i.e., for radial functions φ we have

$$\mathcal{F}_d \left[\mathcal{F}_d \varphi \right] = \varphi.$$

One of the most celebrated results on positive definite functions is their characterization in terms of Fourier transforms established by Bochner in 1932 (for $d = 1$) and 1933 (for general d).

Theorem (Bochner)

A (complex-valued) function $\Phi \in C(\mathbb{R}^d)$ is positive definite on $\mathbb{R}^d$ if and only if it is the Fourier transform of a finite non-negative Borel measure μ on $\mathbb{R}^d$, i.e.,

$$\Phi(\boldsymbol{x}) = \hat{\mu}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^d}} \int_{\mathbb{R}^d} e^{-i\boldsymbol{x}\cdot\boldsymbol{t}} d\mu(\boldsymbol{t}), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

Remark

There are many proofs of this theorem.

Bochner's original proof can be found in [Boc33]. Other proofs can be found, e.g., in [Cup75] or [GV64].

A proof using the Riesz representation theorem to interpret the Borel measure as a distribution, and then take advantage of distributional Fourier transforms can be found in [Wen05].

We will prove only the one (easy) direction that is important for the application to scattered data interpolation.

Proof (easy direction).

We assume Φ is the Fourier transform of a finite non-negative Borel measure and show Φ is positive definite.
Thus,

$$
\begin{aligned}
\sum_{j=1}^{N}\sum_{k=1}^{N} c_j \overline{c_k} \Phi(\boldsymbol{x}_j - \boldsymbol{x}_k) &= \frac{1}{\sqrt{(2\pi)^d}} \sum_{j=1}^{N}\sum_{k=1}^{N} \left[c_j \overline{c_k} \int_{\mathbb{R}^d} e^{-i(\boldsymbol{x}_j - \boldsymbol{x}_k)\cdot \boldsymbol{t}} d\mu(\boldsymbol{t}) \right] \\
&= \frac{1}{\sqrt{(2\pi)^d}} \int_{\mathbb{R}^d} \left[\sum_{j=1}^{N} c_j e^{-i\boldsymbol{x}_j \cdot \boldsymbol{t}} \sum_{k=1}^{N} \overline{c_k} e^{i\boldsymbol{x}_k \cdot \boldsymbol{t}} \right] d\mu(\boldsymbol{t}) \\
&= \frac{1}{\sqrt{(2\pi)^d}} \int_{\mathbb{R}^d} \left| \sum_{j=1}^{N} c_j e^{-i\boldsymbol{x}_j \cdot \boldsymbol{t}} \right|^2 d\mu(\boldsymbol{t}) \geq 0.
\end{aligned}
$$

The last inequality holds because of the conditions imposed on the measure μ. $\qquad\square$

Remark

- *Bochner's theorem shows that for any fixed $\boldsymbol{t} \in \mathbb{R}^d$ the function*

$$\Phi_e(\boldsymbol{x}) = K_e(\boldsymbol{x}, \boldsymbol{0}) = e^{i\boldsymbol{x}\cdot\boldsymbol{t}}, \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

  from our earlier example can be considered as the fundamental positive definite function since all other positive definite functions are obtained as (infinite) linear combinations of this function.

- *We also saw that linear combinations of positive definite kernels/functions will again be positive definite. The remarkable content of Bochner's theorem is the fact that indeed all positive definite functions are generated by Φ_e.*

In order to guarantee a well-posed interpolation problem we have to extend (if possible) Bochner's characterization to *strictly* positive definite functions.

We give a sufficient condition for a function to be strictly positive definite on $\mathbb{R}^d$.

Theorem

Let μ be a non-negative finite Borel measure on $\mathbb{R}^d$ whose carrier is a set of nonzero Lebesgue measure. Then the Fourier transform of μ is strictly positive definite on $\mathbb{R}^d$.

Proof.

As for Bochner, but with some extra measure theoretic arguments (see [CL99]). □

The following corollary gives us a way to *construct* strictly positive definite functions.

Corollary

Let f be a continuous non-negative function in $L_1(\mathbb{R}^d)$ which is not identically zero. Then the Fourier transform of f is strictly positive definite on $\mathbb{R}^d$.

A very useful criterion to *check* whether a given function is strictly positive definite is given by

Theorem

Let Φ be a continuous function in $L_1(\mathbb{R}^d)$. Φ is strictly positive definite if and only if Φ is bounded and its Fourier transform is non-negative and not identically equal to zero.

Proof.

Proof of corollary and theorem in [Wen05].   □

Positive Definite Radial Functions

Earlier we characterized (strictly) positive definite functions in terms of multivariate functions Φ.

When working with radial functions $\Phi(\boldsymbol{x}) = \varphi(\|\boldsymbol{x}\|)$ it is convenient to call the univariate function φ a *positive definite radial function*.

This is a small abuse of our terminology for positive definite functions, but commonly done in the literature.

If follows immediately that

Lemma

If $\Phi = \varphi(\|\cdot\|)$ is (strictly) positive definite and radial on $\mathbb{R}^{d_0}$ then Φ is also (strictly) positive definite and radial on $\mathbb{R}^d$ for any $d \leq d_0$.

We now return to integral characterizations and begin with a theorem due to Schoenberg (see, e.g., [Sch38, p.816], or [WW76, p.27]).

Theorem

A continuous function $\varphi : [0, \infty) \to \mathbb{R}$ is positive definite and radial on $\mathbb{R}^d$ if and only if it is the Bessel transform of a finite non-negative Borel measure μ on $[0, \infty)$, i.e.,

$$\varphi(r) = \int_0^\infty \Omega_d(rt) \mathrm{d}\mu(t).$$

Here

$$\Omega_d(r) = \begin{cases} \cos r & \text{for } d = 1, \\ \Gamma\left(\frac{d}{2}\right)\left(\frac{2}{r}\right)^{\frac{d-2}{2}} J_{\frac{d-2}{2}}(r) & \text{for } d \geq 2, \end{cases}$$

and $J_{\frac{d-2}{2}}$ is the classical Bessel function of the first kind of order $\frac{d-2}{2}$.

Remark

- *Now, for any fixed t, we can view the function $\varphi_t(r) = \cos(rt)$ as the fundamental positive definite radial function on $\mathbb{R}$ since any such function is given by an infinite linear combination from $\{\varphi_t\}_t$.*

- *Moreover, for any fixed t and d the functions $\varphi_{d,t}(r) = \Omega_d(rt)$ can be viewed as the fundamental functions that are positive definite and radial on $\mathbb{R}^d$.*

A Fourier transform characterization of strictly positive definite radial functions on $\mathbb{R}^d$ can be found in [Wen05]. This theorem is based on the Fourier transform formula of radial functions given earlier and the check for strictly positive definite functions.

Theorem

A continuous function $\varphi : [0, \infty) \to \mathbb{R}$ such that
$r \mapsto r^{d-1}\varphi(r) \in L_1[0, \infty)$ is strictly positive definite and radial on $\mathbb{R}^d$ if and only if the d-dimensional Fourier transform

$$\mathcal{F}_d\varphi(r) = \frac{1}{\sqrt{r^{d-2}}} \int_0^\infty \varphi(t) t^{\frac{d}{2}} J_{\frac{d-2}{2}}(rt)dt$$

is non-negative and not identically equal to zero.

Above we saw that any function that is (strictly) positive definite and radial on $\mathbb{R}^{d_0}$ is also (strictly) positive definite and radial on $\mathbb{R}^d$ for any $d \leq d_0$.

Therefore, we are interested in those functions which are (strictly) positive definite and radial on $\mathbb{R}^d$ for all d.

The characterization for positive definite functions is from [Sch38, pp. 817–821] and the strictly positive definite case from [Mic86]:

Theorem (Schoenberg)

A continuous function $\varphi : [0, \infty) \to \mathbb{R}$ is strictly positive definite and radial on $\mathbb{R}^d$ for all d if and only if it is of the form

$$\varphi(r) = \int_0^\infty e^{-r^2 t^2} d\mu(t),$$

where μ is a finite non-negative Borel measure on $[0, \infty)$ not concentrated at the origin.

Letting μ be a point evaluation measure concentrated at $t = \varepsilon > 0$ in Schoenberg's theorem shows us that the Gaussian

$$\varphi_\varepsilon(r) = e^{-\varepsilon^2 r^2}$$

is strictly positive definite and radial on $\mathbb{R}^d$ for all d.

Moreover, we can view the Gaussian as the fundamental member of the family of functions that are strictly positive definite and radial on $\mathbb{R}^d$ for all d since these functions are obtained via infinite linear combinations of Gaussians of different scales ε.

By Schoenberg's theorem *all* such functions that are strictly positive definite and radial on $\mathbb{R}^d$ for all d are given as infinite linear combinations of Gaussians.

Schoenberg's theorem employs a finite non-negative Borel measure μ on $[0, \infty)$ such that

$$\varphi(r) = \int_0^\infty e^{-r^2 t^2} d\mu(t).$$

If we want to find a zero r_0 of φ then we have to solve

$$\varphi(r_0) = \int_0^\infty e^{-r_0^2 t^2} d\mu(t) = 0.$$

Since the exponential function is positive and the measure is non-negative, it follows that μ must be the zero measure.

However, then φ is identically equal to zero.

Therefore, a non-trivial function φ that is positive definite and radial on $\mathbb{R}^d$ *for all d* can have no zeros.

The discussion on the previous slide implies in particular that

Theorem

- *There are no oscillatory univariate continuous functions that are strictly positive definite and radial on $\mathbb{R}^d$ for all d.*

- *There are no compactly supported univariate continuous functions that are strictly positive definite and radial on $\mathbb{R}^d$ for all d.*

Remark

This probably explains the confusion that occurred at the presentation of compactly supported radial basis functions by Robert Schaback in Ulvik, Norway, in 1994.

References I

[Boc33] S. Bochner, *Monotone Funktionen, Stieltjes Integrale und harmonische Analyse*, Math. Ann. **108** (1933), 378–410.

[Buh03] M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 2003.

[CL99] E. W. Cheney and W. A. Light, *A Course in Approximation Theory*, Brooks/Cole (Pacific Grove, CA), 1999.

[Cup75] R. Cuppens, *Decomposition of Multivariate Probability*, Academic Press (New York), 1975.

[GV64] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions Vol. 4*, Academic Press (New York), 1964.

[Mic86] C. A. Micchelli, *Interpolation of scattered data: Distance matrices and conditionally positive definite functions*, Constructive Approximation **2** (1986), no. 1, 11–22.

References II

[Rud91] W. Rudin, *Functional Analysis*, 2nd ed., McGraw-Hill (New York), 1991.

[Sch38] I. J. Schoenberg, *Metric spaces and completely monotone functions*, The Annals of Mathematics **39** (1938), no. 4, 811–841.

[Sne72] I. N. Sneddon, *The Use of Integral Transforms*, McGraw-Hill (New York), 1972.

[SS02]   B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, The MIT Press, 2002.

[SW71]   E. M. Stein and G. Weiss, *Introduction to Fourier Analysis in Euclidean Spaces*, Princeton University Press (Princeton), 1971.

[Wen05] H. Wendland, *Scattered Data Approximation*, Cambridge Monographs on Applied and Computational Mathematics, vol. 17, Cambridge University Press, 2005.

[WW76]   J. H. Wells and L. R. Williams, *Embeddings and Extensions in Analysis*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer, Berlin, 1976.