Advances in Computational Mathematics **20:** 263–271, 2004.

Strictly positive definite functions on a real inner product space

Allan Pinkus
Department of Mathematics, Technion, I. I. T., Haifa, 32000, Israel

Received 5 March 2002; accepted 21 June 2002
Communicated by C.A. Micchelli

If $f(t) = \sum_{k=0}^{\infty} a_k t^k$ converges for all $t \in \mathbb{R}$ with all coefficients $a_k \geqslant 0$, then the function $f(\langle \mathbf{x}, \mathbf{y} \rangle)$ is positive definite on $H \times H$ for any inner product space H. Set $K = \{k\colon a_k > 0\}$. We show that $f(\langle \mathbf{x}, \mathbf{y} \rangle)$ is strictly positive definite if and only if K contains the index 0 plus an infinite number of even integers and an infinite number of odd integers.

Keywords: strictly positive definite, inner product space

AMS subject classification: 42A82, 41A05

1. Introduction

Let H be any real inner product space, $\langle \cdot, \cdot \rangle$ the inner product thereon, and B a subset of H. We will say that a function $f : \mathbb{R} \to \mathbb{R}$ is *positive definite* on B if for any choice of $\mathbf{x}^1, \ldots, \mathbf{x}^n$ in B and any n, the matrix

$$\left(f\left(\left\langle \mathbf{x}^i, \mathbf{x}^j \right\rangle \right) \right)_{i,j=1}^{n}$$

is positive definite. Equivalently, for all $\mathbf{x}^1, \ldots, \mathbf{x}^n$ in B and $c_1, \ldots, c_n \in \mathbb{R}$

$$\sum_{i,j=1}^{n} c_i f\left(\left\langle \mathbf{x}^i, \mathbf{x}^j \right\rangle \right) c_j \geqslant 0.$$

We say that a function $f : \mathbb{R} \to \mathbb{R}$ is *strictly positive definite* on B if for any choice of distinct $\mathbf{x}^1, \ldots, \mathbf{x}^n$ in B and any n, the matrix

$$\left(f\left(\left\langle \mathbf{x}^i, \mathbf{x}^j \right\rangle \right) \right)_{i,j=1}^{n}$$

is strictly positive definite. Equivalently, for $\mathbf{x}^i$ as above, and every $c_1, \ldots, c_n \in \mathbb{R}$, not all zero,

$$\sum_{i,j=1}^{n} c_i f\left(\left\langle \mathbf{x}^i, \mathbf{x}^j \right\rangle \right) c_j > 0.$$

(This is not the usual definition of a (strictly) positive definite function which classically considers $f(\|\mathbf{x}-\mathbf{y}\|)$ rather than $f(\langle\mathbf{x},\mathbf{y}\rangle)$.)

We know that for any $\mathbf{x}^1,\dots,\mathbf{x}^n$ in B the matrix

$$\left(\langle\mathbf{x}^i,\mathbf{x}^j\rangle\right)_{i,j=1}^n$$

is itself positive definite. That is, given $c_1,\dots,c_n\in\mathbb{R}$

$$\sum_{i,j=1}^n c_i\langle\mathbf{x}^i,\mathbf{x}^j\rangle c_j=\sum_{i,j=1}^n\langle c_i\mathbf{x}^i,c_j\mathbf{x}^j\rangle=\left\|\sum_{i=1}^n c_i\mathbf{x}^i\right\|^2\geqslant 0.$$

It is natural to ask for conditions on a function $f:\mathbb{R}\to\mathbb{R}$ such that

$$f\bigl(\langle\mathbf{x},\mathbf{y}\rangle\bigr)$$

is positive definite or, additionally, strictly positive definite on a particular B.

It is known (see, e.g., [7; 1, p. 159; 4]) that if for all positive definite matrices $(a_{ij})_{i,j=1}^n$ (and for all n) the matrix $(f(a_{ij}))_{i,j=1}^n$ is positive definite, then f is necessarily of the form

$$f(t)=\sum_{k=0}^\infty a_kt^k \tag{1.1}$$

with $a_k\geqslant 0$ for all k, and where the series converges for all $t\in\mathbb{R}$ (f is real entire, and absolutely monotone on $\mathbb{R}_+$). Thus if we demand that f be positive definite on H where H is infinite-dimensional, or where $H=\mathbb{R}^d$ with the usual inner product

$$\langle\mathbf{x},\mathbf{y}\rangle=\sum_{i=1}^d x_iy_i$$

and for **all** d, then f is necessarily of the form (1.1). However there are H for which condition (1.1) is not necessary in order that f be positive definite on H. For example, if $H=\mathbb{R}$ with $\langle x,y\rangle=xy$ ($x,y\in\mathbb{R}$), then the functions $f(t)=|t|^p$ and $f(t)=(\operatorname{sgn}t)|t|^p$ give rise to positive definite functions for any $p>0$. To the best of my knowledge, there is no complete characterization of positive definite functions on $\mathbb{R}^d$, for any fixed d.

In this paper we will always assume that f has the form of (1.1). Our main result will be that among all such functions f is strictly positive definite on H if and only if the set

$$K=\{k\colon a_k>0\} \tag{1.2}$$

contains the index 0 plus an infinite number of even integers and an infinite number of odd integers. This result is in fact valid for any B in any real inner product space H which contains an interval symmetric about the origin.

The motivation for considering this question comes partially from certain papers in Learning Theory, see, e.g., [2,8–11]. Positive definite kernels

$$R(\mathbf{x},\mathbf{y}) = f\big(\langle \mathbf{x},\mathbf{y}\rangle\big)$$

give rise to reproducing kernel Hilbert spaces (RKHS). That is, with respect to some suitably chosen inner products $(\cdot\,,\cdot)$ there is defined the reproducing space of functions g defined on B satisfying

$$g(\mathbf{y}) = \big(g(\cdot),\, f\big(\langle \cdot\,,\mathbf{y}\rangle\big)\big)$$

for $\mathbf{y} \in B$. If $f(\langle \mathbf{x},\mathbf{y}\rangle)$ is positive definite, but not strictly positive definite, on B, then

$$\sum_{i,j=1}^{n} c_i f\big(\langle \mathbf{x}^i,\mathbf{x}^j\rangle\big)c_j = 0$$

for some distinct $\mathbf{x}^1,\dots,\mathbf{x}^n \in B$ and $c_1,\dots,c_n$ not all zero. Moreover in this case

$$0 = \sum_{i,j=1}^{n} c_i f\big(\langle \mathbf{x}^i,\mathbf{x}^j\rangle\big)c_j = \left(\sum_{i=1}^{n} c_i f\big(\langle \cdot\,,\mathbf{x}^i\rangle\big), \sum_{j=1}^{n} c_j f\big(\langle \cdot\,,\mathbf{x}^j\rangle\big)\right)$$
$$= \left\| \sum_{i=1}^{n} c_i f\big(\langle \cdot\,,\mathbf{x}^i\rangle\big)\right\|$$

which immediately implies that

$$\sum_{i=1}^{n} c_i g\big(\mathbf{x}^i\big) = 0$$

for every g in this RKHS. This is generally an undesirable property, especially if one wishes to use this RKHS as an approximating set. Thus it is preferable that the kernel be strictly positive definite.

The approximating property just alluded to may have many forms. One possible form (see [9,10] for something similar) is to demand that

$$\operatorname{span}\big\{f\big(\langle \cdot\,,\mathbf{y}\rangle\big):\ \mathbf{y} \in B\big\}$$

be dense in $C(B)$ where B is any compact subset of H. If $H = \mathbb{R}^d$ for any fixed d, and f is positive definite, then the property of $f(\langle \mathbf{x},\mathbf{y}\rangle)$ being strictly positive definite is of course necessary for density. However it is not sufficient. Assuming B contains a neighborhood of the origin and f is of the form (1.1), then the necessary and sufficient condition for density, as above, is that K contains 0, an infinite set of even (positive) indices $\{m_i\}$, and an infinite set of odd indices $\{n_i\}$ such that

$$\sum_{i=1}^{\infty} \frac{1}{m_i} = \sum_{i=1}^{\infty} \frac{1}{n_i} = \infty.$$

This is a Müntz condition. This result may be essentially found in [3]. Moreover, this result has nothing to do with positive definite functions. This density property holds for every real entire function f where K is now the set of integers associated with nonzero coefficients, i.e., the set of k for which $a_k \neq 0$ (and the a_k may be negative). And of course there are other nonentire and nonpositive definite f for which the density property also holds.

2. Main result

The main result of this paper is the following:

Theorem. Assume

$$f(t) = \sum_{k=0}^{\infty} a_k t^k$$

with $a_k \geqslant 0$, all k, is convergent for all t, and

$$K = \{k\colon\ a_k > 0\}.$$

Let H be any real inner product space and let B be a subset of H containing $\{\alpha\mathbf{z}\colon |\alpha| \leqslant c\}$ for some $c > 0$ and $\mathbf{z} \in H$, $\|\mathbf{z}\| = 1$. Then f is strictly positive definite on B if and only if K contains the index 0 plus an infinite number of even integers and an infinite number of odd integers.

Proof. We will first prove the necessity in the case $H = \mathbb{R}$ and $\langle x, y\rangle = xy$. We assume $[-c, c] \subseteq B$ for some $c > 0$ and let $x_1, \ldots, x_n$ be n distinct real values in B. The matrix

$$\big(f(x_i x_j)\big)_{i,j=1}^{n}$$

is strictly positive definite if and only if

$$\sum_{i,j=1}^{n} c_i f(x_i x_j) c_j > 0 \tag{2.1}$$

for every choice of $c_1, \ldots, c_n$ not all zero. As

$$f(t) = \sum_{k=0}^{\infty} a_k t^k$$

is convergent for all t, we may rewrite (2.1) as

$$\sum_{k=0}^{\infty} a_k \sum_{i,j=1}^{n} c_i (x_i x_j)^k c_j = \sum_{k=0}^{\infty} a_k \left(\sum_{i=1}^{n} c_i x_i^k\right)^2.$$

Since the $a_k \geqslant 0$ for all k, we have

$$\sum_{k=0}^{\infty} a_k \left(\sum_{i=1}^{n} c_i x_i^k \right)^2 \geqslant 0.$$

Our aim is to determine exact conditions on K implying strict positivity for all choices of $c_1, \ldots, c_n$ not all zero. In other words, we want necessary and sufficient conditions on K such that there exist no choice of distinct points $x_1, \ldots, x_n$ in B and $c_1, \ldots, c_n$ not all zero such that

$$\sum_{i=1}^{n} c_i x_i^k = 0 \tag{2.2}$$

for all $k \in K$.

To see that we must have $0 \in K$, let $x_1 = 0$, $c_1 \neq 0$ and $n = 1$. Then

$$c_1 f(0) c_1 = a_0 c_1^2 > 0$$

implies that $a_0 > 0$, i.e., $0 \in K$.

Assume K contains only a finite number, say m, of even indices. Then given any distinct strictly positive $x_1, \ldots, x_{m+1}$ in $(0, c]$ there exist $c_1, \ldots, c_{m+1}$, not all zero, such that

$$\sum_{i=1}^{m+1} c_i x_i^k = 0$$

for each $k \in K$ which is even. Thus

$$\sum_{i=1}^{m+1} c_i x_i^k + \sum_{i=1}^{m+1} c_i (-x_i)^k = 0$$

for each $k \in K$ which is even, and every k odd. In other words we have found a sum of the form (2.2) with all $\pm x_i \in B$ which vanishes for all $k \in K$. Similarly, if K contains only a finite number, say m, of odd indices, then given distinct strictly positive $x_1, \ldots, x_{m+1}$ in $(0, c]$ there exist $c_1, \ldots, c_{m+1}$, not all zero, such that

$$\sum_{i=1}^{m+1} c_i x_i^k = 0$$

for each $k \in K$ which is odd. Thus

$$\sum_{i=1}^{m+1} c_i x_i^k - \sum_{i=1}^{m+1} c_i (-x_i)^k = 0$$

for each $k \in K$ which is odd, and every k even. In other words, we have found a sum of the form (2.2) with $\pm x_i \in B$ which vanishes for all $k \in K$. This proves that if $f(xy)$

is strictly positive definite on B then K must contain the index 0 plus an infinite number of even integers and an infinite number of odd integers.

Now assume we have any real inner product space H, and $B \subseteq H$ as in the statement of the theorem. We show that the necessity follows easily from the case $H = \mathbb{R}$. Let us assume that the matrices

$$\left(f\left(\langle \mathbf{x}^i, \mathbf{x}^j \rangle\right)\right)_{i,j=1}^n$$

are strictly positive definite for all choices of distinct $\mathbf{x}^1, \ldots, \mathbf{x}^n \in B$ and all n. Then for any distinct $x_1, \ldots, x_n \in [-c, c]$, and all n, the matrices

$$\left(f\left(\langle x_i \mathbf{z}, x_j \mathbf{z} \rangle\right)\right)_{i,j=1}^n = \left(f(x_i x_j)\right)_{i,j=1}^n$$

are strictly positive definite. (Recall that $\langle \mathbf{z}, \mathbf{z} \rangle = 1$.) This is the previous case with $B = [-c, c]$. Thus necessarily K must contain the index 0 plus an infinite number of even integers and an infinite number of odd integers.

We now prove the sufficiency in the case $H = \mathbb{R}$. We assume that $0 \in K$ and K contains an infinite number of even integers and an infinite number of odd integers. We will show that f is strictly positive definite on all of $\mathbb{R}$, and thus on every $B \subseteq \mathbb{R}$. From the above it suffices to show that there cannot exist distinct $x_1, \ldots, x_n$, and $c_1, \ldots, c_n$ not all zero, some n, such that

$$\sum_{i=1}^n c_i x_i^k = 0$$

for each $k \in K$. It is easily shown that since $0 \in K$ it in fact suffices to show that there cannot exist distinct nonzero $x_1, \ldots, x_n$, and nonzero $c_1, \ldots, c_n$, some n, such that

$$\sum_{i=1}^n c_i x_i^k = 0 \tag{2.3}$$

for each $k \in K' = K \backslash \{0\}$.

Assume this is not the case, and such c_i and x_i as in (2.3) do exist. Let

$$\lambda = \max\left\{|x_i|\colon\ i = 1, \ldots, n\right\},$$

and

$$y_i = \frac{x_i}{\lambda}.$$

Thus

$$\sum_{i=1}^n c_i y_i^k = 0$$

for an infinite number of (even and odd) integers k. Given any $\varepsilon > 0$ there exists an N such that for $k > N$

$$\left|\sum_{|y_i|<1} c_i y_i^k\right| < \varepsilon.$$

Thus

$$\left|\sum_{|y_i|=1} c_i y_i^k\right| < \varepsilon$$

for an infinite number of even and odd integers k. The set $\{i\colon |y_i| = 1\}$ contains either one or two values. If it contains one value i, then $|c_i| < \varepsilon$ for every $\varepsilon > 0$, a contradiction. If it contains two values i and j, then

$$|c_i \pm c_j| < \varepsilon$$

implying that

$$|c_i|, |c_j| < \varepsilon$$

for every $\varepsilon > 0$, which is again a contradiction. This contradiction to our assumption proves sufficiency in the case $H = \mathbb{R}$.

It remains to prove the sufficiency in the case of any real inner product space H. Our proof is based on the following result. Let $\mathbf{x}^1, \dots, \mathbf{x}^n$ be distinct points in H. Then there exist distinct real numbers $b_1, \dots, b_n$ and a positive definite matrix $(m_{ij})_{i,j=1}^n$ such that

$$\langle \mathbf{x}^i, \mathbf{x}^j \rangle = b_i b_j + m_{ij}.$$

We defer the proof of this result. As the $(m_{ij})_{i,j=1}^n$ is a positive definite matrix, we have

$$\sum_{i,j=1}^n c_i \langle \mathbf{x}^i, \mathbf{x}^j \rangle c_j = \sum_{i,j=1}^n c_i b_i b_j c_j + \sum_{i,j=1}^n c_i m_{ij} c_j \geqslant \sum_{i,j=1}^n c_i b_i b_j c_j.$$

Similarly, it is well known that

$$\sum_{i,j=1}^n c_i \langle \mathbf{x}^i, \mathbf{x}^j \rangle^k c_j \geqslant \sum_{i,j=1}^n c_i (b_i b_j)^k c_j.$$

(This is a consequence of inequalities between the Schur (Hadamard) product of two positive definite matrices. See, e.g., [5, p. 310].) Thus

$$\begin{aligned}\sum_{i,j=1}^n c_i f\left(\langle \mathbf{x}^i, \mathbf{x}^j \rangle\right) c_j &= \sum_{k=0}^\infty a_k \sum_{i,j=1}^n c_i \langle \mathbf{x}^i, \mathbf{x}^j \rangle^k c_j \\ &\geqslant \sum_{k=0}^\infty a_k \sum_{i,j=1}^n c_i (b_i b_j)^k c_j = \sum_{k=0}^\infty a_k \left(\sum_{i=1}^n c_i b_i^k\right)^2.\end{aligned}$$

However as the b_i are distinct real values and K contains 0 plus an infinite number of even integers and an infinite number of odd integers, we know that the last term in the above sequence of inequalities is strictly positive for all choices of $c_1, \ldots, c_n$ not all zero. This proves the sufficiency. $\square$

It remains to prove the following result.

Proposition. Let H be any real inner product space, and $\mathbf{x}^1, \ldots, \mathbf{x}^n$ be n distinct points in H. Then there exist distinct real numbers $b_1, \ldots, b_n$ and a positive definite matrix $(m_{ij})_{i,j=1}^n$ such that

$$\langle \mathbf{x}^i, \mathbf{x}^j \rangle = b_i b_j + m_{ij}.$$

Proof. Set

$$a_{ij} = \langle \mathbf{x}^i, \mathbf{x}^j \rangle.$$

We first claim that as the $\mathbf{x}^1, \ldots, \mathbf{x}^n$ are distinct points in H, the positive definite matrix $A = (a_{ij})_{i,j=1}^n$ has no two identical rows (or columns). Assume this is not the case. Then for some $i, j \in \{1, \ldots, n\}$ we have

$$a_{ii} = a_{ij} = a_{ji} = a_{jj}$$

which is the same as

$$\langle \mathbf{x}^i, \mathbf{x}^i \rangle = \langle \mathbf{x}^i, \mathbf{x}^j \rangle = \langle \mathbf{x}^j, \mathbf{x}^j \rangle.$$

It is now easily checked that

$$\left\| \langle \mathbf{x}^j, \mathbf{x}^j \rangle \mathbf{x}^i - \langle \mathbf{x}^i, \mathbf{x}^j \rangle \mathbf{x}^j \right\| = 0$$

which implies that

$$\mathbf{x}^i = \mathbf{x}^j$$

contradicting our assumption.

The remaining steps of the proof may be found as an exercise in [6, p. 287]. For completeness we record it here.

Since A is a positive definite matrix it may be decomposed as

$$A = C^{\mathrm{T}} C$$

where $C = (c_{kj})_{k=1}^{m}{}_{j=1}^{n}$ is an $m \times n$ matrix and $\operatorname{rank} A = m$. Thus

$$a_{ij} = \sum_{k=1}^{m} c_{ki} c_{kj}, \qquad i, j = 1, \ldots, n.$$

As no two rows of A are identical, it follows that no two columns of C are identical. There therefore exists a $\mathbf{v} \in \mathbb{R}^m$, $\|\mathbf{v}\| = 1$, for which $\mathbf{v}C = \mathbf{b}$ with $\mathbf{b} = (b_1, \ldots, b_n)$ where the $\{b_i\}_{i=1}^n$ are distinct.

Now let V be any $m \times m$ orthogonal matrix ($V^{\mathrm{T}}V = I$) whose first row is the above $\mathbf{v}$. Let U be the $m \times m$ matrix whose first row is the above $\mathbf{v}$ and all its other entries are zero. Then

$$A = C^{\mathrm{T}}C = C^{\mathrm{T}}V^{\mathrm{T}}VC = C^{\mathrm{T}}U^{\mathrm{T}}UC + C^{\mathrm{T}}(V-U)^{\mathrm{T}}(V-U)C$$

since $U^{\mathrm{T}}(V-U) = (V-U)^{\mathrm{T}}U = 0$. From the above, it follows that $(C^{\mathrm{T}}U^{\mathrm{T}}UC)_{ij} = b_i b_j$ and $M = C^{\mathrm{T}}(V-U)^{\mathrm{T}}(V-U)C$ is positive definite. This completes the proof of the proposition. $\square$

Remark. If B is bounded, then the above theorem will hold for every $f(t) = \sum_{k=0}^{\infty} a_k t^k$, $a_k \geqslant 0$, which converges in the ball of radius $\sup\{\|\mathbf{x}\|\colon \mathbf{x} \in B\}$.

References

[1] C. Berg, J.P.R. Christensen and P. Ressel, *Harmonic Analysis on Semigroups* (Springer, New York, 1984).
[2] C.J.C. Burges and D.J. Crisp, Uniqueness of the SVM solution, in: *NIPS (Neural Information Processing Systems)*, Vol. 12, eds. S.A. Solla, T.K. Leen and K.-R. Müller (2000) pp. 223–229.
[3] W. Dahmen and C.A. Micchelli, Some remarks on ridge functions, Approx. Theory Appl. 3 (1987) 139–143.
[4] C.H. FitzGerald, C.A. Micchelli and A. Pinkus, Functions that preserve families of positive semidefinite matrices, Linear Algebra Appl. 221 (1995) 83–102.
[5] R. Horn and C.R. Johnson, *Topics in Matrix Analysis* (Cambridge Univ. Press, Cambridge, 1991).
[6] G. Pólya and G. Szegö, *Problems and Theorems in Analysis*, Vol. II (Springer, Berlin, 1976).
[7] I.J. Schoenberg, Positive definite functions on spheres, Duke Math. J. 9 (1942) 96–108.
[8] A.J. Smola, B. Schölkopf and K.-R. Müller, The connection between regularization operators and support vector kernels, Neural Networks 11 (1998) 637–649.
[9] I. Steinwart, On the influence of the kernel on the consistency of support vector machines, J. Machine Learning Res. 2 (2001) 67–93.
[10] I. Steinwart, Support vector machines are universally consistent, Preprint.
[11] V.N. Vapnik, *The Nature of Statistical Learning Theory* (Springer, New York, 1995).