Aronszajn’s theorem

Jean-Philippe Vert

We detail the proof of Aronszajn’s theorem which shows the equivalence between being positive definiteness (p.d.) and being a reproducing kernel (r.k.), as shown by ?. Recall that p.d. and r.k. kernels are defined as follows:

**Definition 1.** Let \( X \) be a set. A function \( K : X \times X \to \mathbb{R} \) is called a positive definite kernel on \( X \) iff it is symmetric, that is, \( K(x, x') = K(x', x) \) for any two objects \( x, x' \in X \), and positive definite, that is,
\[
\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j K(x_i, x_j) \geq 0
\]
for any \( n > 0 \), any choice of \( n \) points \( x_1, \cdots, x_n \in X \), and any choice of real numbers \( c_1, \cdots, c_n \in \mathbb{R} \).

**Definition 2.** Let \( X \) be a set and \( H \subset \mathbb{R}^X \) be a class of functions forming a Hilbert space with inner product \( \langle \cdot, \cdot \rangle_H \). The function \( K : X^2 \to \mathbb{R} \) is called a reproducing kernel (r.k.) of \( H \) if

1. \( H \) contains all functions of the form
\[
\forall x \in X, \quad K_x : t \mapsto K(x, t) .
\]

2. For every \( x \in X \) and \( f \in H \) the reproducing property holds:
\[
f(x) = \langle f, K_x \rangle_H .
\]

If a r.k. exists, then \( H \) is called a reproducing kernel Hilbert space (RKHS).

Aronszajn’s theorem now states that:

**Theorem 1.** For any set \( X \), a function \( K : X \times X \) is positive definite if and only if it is a reproducing kernel.

**Proof.** Let us first assume that \( K \) is the r.k. of an RKHS \( H \). Then it is symmetric because, for any \((x, y) \in X^2\), we can use the symmetry of the inner product in \( H \) to get:
\[
K(x, y) = \langle K_x, K_y \rangle_H = \langle K_y, K_x \rangle_H = K(y, x) .
\]

Moreover, for any \( N \in \mathbb{N}, (x_1, x_2, \ldots, x_N) \in X^N \), and \( (a_1, a_2, \ldots, a_N) \in \mathbb{R}^N \):
\[
\sum_{i, j=1}^{N} a_i a_j K(x_i, x_j) = \sum_{i, j=1}^{N} a_i a_j \langle K_{x_i}, K_{x_j} \rangle_H
= \| \sum_{i=1}^{N} a_i K_{x_i} \|_{H}^2
\geq 0 .
\]
\( K \) is therefore p.d. Conversely, let us now suppose that \( K \) is p.d. In order to build a RKHS having \( K \) as r.k., we start by considering the vector space \( \mathcal{H}_0 \subset \mathbb{R}^X \) spanned by the functions \( \{ K_x \}_{x \in X} \). For any \( f, g \in \mathcal{H}_0 \), given by:

\[
 f = \sum_{i=1}^{m} a_i K_{x_i}, \quad g = \sum_{j=1}^{n} b_j K_{y_j},
\]

let us define the operation:

\[
 \langle f, g \rangle_{\mathcal{H}_0} := \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j K(x_i, y_j).
\]

We note that \( \langle f, g \rangle_{\mathcal{H}_0} \) does not depend on the expansion of \( f \) and \( g \), because:

\[
 \langle f, g \rangle_{\mathcal{H}_0} = \sum_{i=1}^{m} a_i g(x_i) = \sum_{j=1}^{n} b_j f(y_j).
\]

This also shows that \( \langle \, . \, \rangle_{\mathcal{H}_0} \) is a symmetric bilinear form. Moreover, for any \( x \in X \) and \( f \in \mathcal{H}_0 \):

\[
 \langle f, K_x \rangle_{\mathcal{H}_0} = f(x).
\]

Now, \( K \) being assumed to be p.d., we also have:

\[
 \| f \|_{\mathcal{H}_0}^2 = \sum_{i,j=1}^{m} a_i a_j K(x_i, x_j) \geq 0.
\]

In particular Cauchy-Schwarz inequality is valid with \( \langle \, \cdot \, \rangle_{\mathcal{H}_0} \). We deduce that \( \forall x \in X \):

\[
 |f(x)| = |\langle f, K_x \rangle_{\mathcal{H}_0}| \leq \| f \|_{\mathcal{H}_0} \| K(x, x) \|^{1/2},
\]

therefore \( \| f \|_{\mathcal{H}_0} = 0 \implies f = 0 \). In other words, \( \mathcal{H}_0 \) is a pre-Hilbert space endowed with the inner product \( \langle \, \cdot \, \rangle_{\mathcal{H}_0} \).

At this step, we have built a pre-Hilbert space which has all the properties of a RKHS for \( K \), except for the completeness. We now need to extend \( \mathcal{H}_0 \) to make it complete. For that purpose, let us first note that for any Cauchy sequence \( (f_n)_{n \geq 0} \) in \( \mathcal{H}_0, \langle \, \cdot \, \rangle_{\mathcal{H}_0} \), it holds that:

\[
 \forall (x, m, n) \in X \times \mathbb{N}^2, \quad |f_m(x) - f_n(x)| \leq \| f_m - g_n \|_{\mathcal{H}_0} \| K(x, x) \|^{1/2}.
\]

This shows that for any \( x \) the sequence \( (f_n(x))_{n \geq 0} \) is Cauchy in \( \mathbb{R} \) and has therefore a limit. Let us now consider \( \mathcal{H} \subset \mathbb{R}^X \) to be the set of functions \( f : X \to \mathbb{R} \) which are pointwise limits of Cauchy sequences in \( \mathcal{H}_0 \), i.e., if \( (f_n) \) is a Cauchy sequence in \( \mathcal{H}_0 \), then \( f(x) = \lim_{n \to \infty} f_n(x) \). We can observe that \( \mathcal{H}_0 \subset \mathcal{H} \).

Indeed, for any \( f \in \mathcal{H}_0 \), it suffices to take the constant function \( f_n = f \) for any \( n \geq 0 \) to obtain a Cauchy sequence in \( \mathcal{H}_0 \) which converges pointwise to \( f \). We shall now define an inner product on \( \mathcal{H} \), and show that \( \mathcal{H} \) endowed with that inner product it is a RKHS with reproducing kernel \( K \).

For that purpose, let us first show a useful property of Cauchy sequences in \( \mathcal{H}_0 \).

**Lemma 1.** Any Cauchy sequence \( (f_n)_{n \in \mathbb{N}} \) in \( \mathcal{H}_0 \) which converges pointwise to 0 satisfies:

\[
 \lim_{n \to \infty} \| f_n \|_{\mathcal{H}_0} = 0.
\]

Indeed, let \( (f_n) \) be a Cauchy sequence in \( \mathcal{H}_0 \). Any Cauchy sequence being bounded, let \( B > \| f_n \| \) for any \( n \in \mathbb{N} \). For any \( \varepsilon > 0 \), let \( N \in \mathbb{N} \) be such that, for any \( n > N \), \( \| f_n - f_N \| < \varepsilon / B \). The function \( f_N \in \mathcal{H}_0 \) can be expanded as:

\[
 f_N(x) = \sum_{i=1}^{p} \alpha_i K(x_i, x),
\]
for some $p \in \mathbb{N}, \alpha_1, \ldots, \alpha_p \in \mathbb{R}$ and $x_1, \ldots, x_p \in X$. We then get, for any $n > N$:

\[
\|f_n\|^2_{\mathcal{H}_0} = \langle f_n - f_N, f_n \rangle_{\mathcal{H}_0} + \langle f_N, f_n \rangle_{\mathcal{H}_0} \\
\leq \varepsilon + \sum_{i=1}^p \alpha_i f_n(x_i).
\]

Since $f_n(x_i)$ converges to 0 for $i = 1, \ldots, p$, we obtain that $\|f_n\|_{\mathcal{H}_0} < 2\varepsilon$ for $n$ large enough, i.e., $\|f_n\|_{\mathcal{H}_0}$ converges to 0. This proves Lemma 1.

Coming back to the proof of Theorem 1, let us consider two Cauchy sequences $(f_n)$ and $(g_n)$ in $\mathcal{H}_0$. These sequences define two functions $f$ and $g$ in $\mathcal{H}$ as their pointwise limits. Let us first show that the inner product $\langle f_n, g_n \rangle_{\mathcal{H}_0}$ converges. For that purpose, we note using the Cauchy-Schwarz inequality that, for any $n, m \in \mathbb{N}$

\[
\left| \langle f_n, g_n \rangle_{\mathcal{H}_0} - \langle f_m, g_m \rangle_{\mathcal{H}_0} \right| = \left| \langle f_n, g_n - g_m \rangle_{\mathcal{H}_0} + \langle f_m, g_n - g_m \rangle_{\mathcal{H}_0} \right| \\
\leq \|f_n\|_{\mathcal{H}_0} \|g_n - g_m\|_{\mathcal{H}_0} + \|f_m - f_n\|_{\mathcal{H}_0} \|g_n - g_m\|_{\mathcal{H}_0}.
\]

Since each Cauchy sequence is bounded in norm, we obtain that $(\langle f_n, g_n \rangle_{\mathcal{H}_0})_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. We have thus shown that the inner product $\langle f_n, g_n \rangle_{\mathcal{H}_0}$ converges. Let us now show that the limit value only depends on the pointwise limits $f$ and $g$. For that purpose, let $(f'_n)$ and $(g'_n)$ be two other Cauchy sequences in $\mathcal{H}_0$ which also converge pointwisely to $f$ and $g$, respectively. Then the sequence $(f_n - f'_n)$ (resp. $(g_n - g'_n)$) is a Cauchy sequence in $\mathcal{H}_0$ which converges pointwisely to 0, and from Lemma 1 we obtain that $\lim_{n \to +\infty} \|f_n - f'_n\|_{\mathcal{H}_0} = 0$ (resp. $\lim_{n \to +\infty} \|g_n - g'_n\|_{\mathcal{H}_0} = 0$). Now we observe that:

\[
\left| \langle f_n, g_n \rangle_{\mathcal{H}_0} - \langle f'_n, g'_n \rangle_{\mathcal{H}_0} \right| = \left| \langle f_n, g_n - g'_n \rangle_{\mathcal{H}_0} + \langle f_n - f'_n, g'_n \rangle_{\mathcal{H}_0} \right| \\
\leq \|f_n\|_{\mathcal{H}_0} \|g_n - g'_n\|_{\mathcal{H}_0} + \|f_n - f'_n\|_{\mathcal{H}_0} \|g'_n\|_{\mathcal{H}_0}.
\]

Both Cauchy sequences being upper bounded in norm, this shows that $\langle f_n, g_n \rangle_{\mathcal{H}_0}$ and $\langle f'_n, g'_n \rangle_{\mathcal{H}_0}$, have the same limit, which only depends on $f$ and $g$. This allows to define formally, for any $f, g \in \mathcal{H}$ defined as pointwise limits of Cauchy sequences $(f_n)$ and $(g_n)$ in $\mathcal{H}_0$:

\[
\langle f, g \rangle_{\mathcal{H}} = \lim_{n \to +\infty} \langle f_n, g_n \rangle_{\mathcal{H}_0}.
\]

It is easy to see that $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ is a positive bilinear form, using the same properties of $\langle \cdot, \cdot \rangle_{\mathcal{H}_0}$. Let now a function $f \in \mathcal{H}$ such that $\|f\|_{\mathcal{H}}^2 = \langle f, f \rangle_{\mathcal{H}} = 0$. By definition $f$ is a pointwise limit of a Cauchy sequence $(f_n)$ in $\mathcal{H}_0$, and $0 = \|f\|_{\mathcal{H}} = \lim_{n \to +\infty} \|f_n\|_{\mathcal{H}_0}$. We then obtain, for any $x \in X$,

\[
|f(x)| = \lim_{n \to +\infty} |f_n(x)| \\
= \lim_{n \to +\infty} |\langle f_n, Kx \rangle_{\mathcal{H}_0}| \\
\leq K(x, x)^{\frac{1}{2}} \times \lim_{n \to +\infty} \|f_n\|_{\mathcal{H}_0} \\
= 0,
\]

showing that $f = 0$. This shows that $\mathcal{H}$ is a pre-Hilbert space with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. Moreover, for any $f \in \mathcal{H}$ defined as the pointwise limit of a Cauchy function $(f_n)$ in $\mathcal{H}_0$, we note that $f_n \in \mathcal{H}$ for any $n \in \mathbb{N}$, and that

\[
\lim_{n \to +\infty} \|f - f_n\|_{\mathcal{H}} = \lim_{n \to +\infty} \lim_{p \to +\infty} \|f_p - f_n\|_{\mathcal{H}_0} = 0.
\]  

This shows in particular that $\mathcal{H}_0$ is dense in $\mathcal{H}$, with respect to the topology defined by the metric $\|\cdot\|_{\mathcal{H}}$. Let us now show the completeness of $\mathcal{H}$. For that purpose, let $(f_n)$ be a Cauchy sequence in $\mathcal{H}$. By
density of \( \mathcal{H}_0 \) in \( \mathcal{H} \), for each \( n \in \mathbb{N} \) we can define a function \( f'_n \in \mathcal{H}_0 \) such that \( \lim_{n \to +\infty} \| f_n - f'_n \|_{\mathcal{H}} = 0 \). For every \( \varepsilon > 0 \), let \( N \in \mathbb{N} \) be such that, for every \( n, m > N \), \( \| f_n - f_m \|_{\mathcal{H}} < \varepsilon/3 \) and \( \| f_n - f'_n \|_{\mathcal{H}} < \varepsilon/3 \).

Using the fact that the norms \( \| \cdot \|_{\mathcal{H}_0} \) and \( \| \cdot \|_{\mathcal{H}} \) coincide on \( \mathcal{H}_0 \), we then obtain, for any \( n, m > N \):

\[
\| f'_n - f'_m \|_{\mathcal{H}_0} = \| f'_n - f'_m \|_{\mathcal{H}} \\
\leq \| f'_n - f_n \|_{\mathcal{H}} + \| f_n - f_m \|_{\mathcal{H}} + \| f_m - f'_m \|_{\mathcal{H}} \\
\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
\leq \varepsilon.
\]

(4)

This shows that \( (f'_n) \) is a Cauchy sequence in \( \mathcal{H}_0 \), which therefore defines a function \( f \in \mathcal{H} \) by pointwise convergence. Moreover this function satisfies, by (3),

\[
\lim_{n \to +\infty} \| f - f'_n \|_{\mathcal{H}} = 0,
\]

and therefore

\[
\lim_{n \to +\infty} \| f - f_n \|_{\mathcal{H}} \leq \lim_{n \to +\infty} \| f - f'_n \|_{\mathcal{H}} + \lim_{n \to +\infty} \| f'_n - f_n \|_{\mathcal{H}} = 0.
\]

This shows that \( f \in \mathcal{H} \) is the limit of the Cauchy sequence \( (f_n) \), and therefore that \( \mathcal{H} \) is complete. \( \mathcal{H} \) is therefore a Hilbert space of functions.

To conclude the proof and show that \( \mathcal{H} \) is a RKHS which admits \( K \) as r.k., we further need to show that the properties (1) and (2) are fulfilled. Condition (1) is immediate since for any \( x \in \mathcal{X} \), by construction, \( K_x \in \mathcal{H}_0 \) and \( \mathcal{H}_0 \subset \mathcal{H} \). To prove (2), let \( x \in \mathcal{X} \) and \( f \in \mathcal{H} \). \( f \) is defined pointwisely as the limit of a Cauchy sequence \( (f_n) \) in \( \mathcal{H}_0 \), and by construction of the inner product in \( \mathcal{H} \) satisfies

\[
f(x) = \lim_{n \to +\infty} f_n(x) \\
= \lim_{n \to +\infty} \langle f_n, K_x \rangle_{\mathcal{H}_0} \\
= \langle f, K_x \rangle_{\mathcal{H}}.
\]

This concludes the proof of Theorem 1.