Aronszajn's theorem

Jean-Philippe Vert

We detail the proof of Aronszajn's theoram which shows the equivalence between being positive definiteness (p.d.) and being a reproducing kernel (r.k.), as shown by (1). Recall that p.d. and r.k. kernels are defined as follows:

Definition 1. *Let $\mathcal{X}$ be a set. A function $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is called a* positive definite kernel *on $\mathcal{X}$ iff it is symmetric, that is, $K(\mathbf{x}, \mathbf{x}') = K(\mathbf{x}', \mathbf{x})$ for any two objects $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, and positive definite, that is,*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0$$

for any $n > 0$, any choice of n points $\mathbf{x}_1, \cdots, \mathbf{x}_n \in \mathcal{X}$, and any choice of real numbers $c_1, \cdots, c_n \in \mathbb{R}$.

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

1. *$\mathcal{H}$ contains all functions of the form*

$$\forall \mathbf{x} \in \mathcal{X}, \quad K_{\mathbf{x}} : \mathbf{t} \mapsto K(\mathbf{x}, \mathbf{t}) \,. \tag{1}$$

2. *For every $\mathbf{x} \in \mathcal{X}$ and $f \in \mathcal{H}$ the emphreproducing property holds:*

$$f(\mathbf{x}) = \langle f, K_{\mathbf{x}} \rangle_{\mathcal{H}} \,. \tag{2}$$

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

Aronszajn's theorem now states that:

Theorem 1. *For any set $\mathcal{X}$, a function $K : \mathcal{X} \times \mathcal{X}$ 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 $\mathcal{H}$. Then it is symmetric because, for any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^2$, we can use the symmetry of the inner product in $\mathcal{H}$ to get:

$$K(\mathbf{x}, \mathbf{y}) = \langle K_{\mathbf{x}}, K_{\mathbf{y}} \rangle_{\mathcal{H}} = \langle K_{\mathbf{y}}, K_{\mathbf{x}} \rangle_{\mathcal{H}} = K(\mathbf{y}, \mathbf{x}) \,.$$

Moreover, for any $N \in \mathbb{N}$,$(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \in \mathcal{X}^N$, and $(a_1, a_2, \ldots, a_N) \in \mathbb{R}^N$:

$$\begin{aligned} \sum_{i,j=1}^{N} a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) &= \sum_{i,j=1}^{N} a_i a_j \left\langle K_{\mathbf{x}_i}, K_{\mathbf{x}_j} \right\rangle_{\mathcal{H}} \\ &= \| \sum_{i=1}^{N} a_i K_{\mathbf{x}_i} \|_{\mathcal{H}}^2 \\ &\geq 0 \,. \end{aligned}$$

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 $\mathcal{H}_0 \subset \mathbb{R}^{\mathcal{X}}$ spanned by the functions $\{K_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{X}}$. For any $f, g \in \mathcal{H}_0$, given by:

$$f = \sum_{i=1}^{m} a_i K_{\mathbf{x}_i}, \quad g = \sum_{j=1}^{n} b_j K_{\mathbf{y}_j},$$

let us define the operation:

$$\langle f, g \rangle_{\mathcal{H}_0} := \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j K(\mathbf{x}_i, \mathbf{y}_j).$$

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

$$\langle f, g \rangle_{\mathcal{H}_0} = \sum_{i=1}^{m} a_i g(\mathbf{x}_i) = \sum_{j=1}^{n} b_j f(\mathbf{y}_j).$$

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

$$\langle f, K_{\mathbf{x}} \rangle_{\mathcal{H}_0} = f(\mathbf{x}).$$

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

$$\| f \|_{\mathcal{H}_0}^2 = \sum_{i,j=1}^{m} a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0.$$

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

$$| f(\mathbf{x}) | = \left| \langle f, K_{\mathbf{x}} \rangle_{\mathcal{H}_0} \right| \leq \| f \|_{\mathcal{H}_0} . K(\mathbf{x}, \mathbf{x})^{\frac{1}{2}},$$

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 $\langle ., . \rangle_{\mathcal{H}_0}$.

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 $\left(\mathcal{H}_0, \langle ., . \rangle_{\mathcal{H}_0}\right)$, it holds that:

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

This shows that for any $\mathbf{x}$ the sequence $(f_n(\mathbf{x}))_{n\geq 0}$ is Cauchy in $\mathbb{R}$ and has therefore a limit. Let us now consider $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$ to be the set of functions $f : \mathcal{X} \to \mathbb{R}$ 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:

$$f_N(x) = \sum_{i=1}^{p} \alpha_i K(\mathbf{x}_i, \mathbf{x}),$$

for some $p \in \mathbb{N}, \alpha_1, \ldots, \alpha_p \in \mathbb{R}$ and $x_1, \ldots, x_p \in \mathcal{X}$. We then get, for any $n > N$:

$$\begin{aligned} \| f_n \|^2_{\mathcal{H}_0} &= \langle f_n - f_N, f_n \rangle_{\mathcal{H}_0} + \langle f_N, f_n \rangle_{\mathcal{H}_0} \\ &\leq \varepsilon + \sum_{i=1}^{p} \alpha_i f_n(\mathbf{x}_i) \,. \end{aligned}$$

Since $f_n(\mathbf{x}_i)$ 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 $n, m \in \mathbb{N}$

$$\begin{aligned} \left| \langle f_n, g_n \rangle_{\mathcal{H}_0} - \langle f_m, g_m \rangle_{\mathcal{H}_0} \right| &= \left| \langle f_n, g_n - g_m \rangle_{\mathcal{H}_0} + \langle f_n - f_m, g_m \rangle_{\mathcal{H}_0} \right| \\ &\leq \| f_n \|_{\mathcal{H}_0} \| g_n - g_m \|_{\mathcal{H}_0} + \| f_n - f_m \|_{\mathcal{H}_0} \| g_m \|_{\mathcal{H}_0} \,. \end{aligned}$$

Since each Cauchy sequence is bounded in norm, we obtain that $\left(\langle f_n, g_n \rangle_{\mathcal{H}_0}\right)_{n \in \mathbb{N}}$ 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. $\lim_{n \to +\infty} \| f_n - f'_n \|_{\mathcal{H}_0} = 0$). Now we observe that:

$$\begin{aligned} \left| \langle f_n, g_n \rangle_{\mathcal{H}_0} - \langle f'_n, g'_n \rangle_{\mathcal{H}_0} \right| &= \left| \langle f_n, g_n - g'_n \rangle_{\mathcal{H}_0} + \langle f_n - f'_n, g_n \rangle_{\mathcal{H}_0} \right| \\ &\leq \| f_n \|_{\mathcal{H}_0} \| g_n - g'_n \|_{\mathcal{H}_0} + \| f_n - f'_n \|_{\mathcal{H}_0} \| g_n \|_{\mathcal{H}_0} \,. \end{aligned}$$

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

$$\begin{aligned} | f(\mathbf{x}) | &= \lim_{n \to +\infty} | f_n(\mathbf{x}) | \\ &= \lim_{n \to +\infty} \left| \langle f_n, K_{\mathbf{x}} \rangle_{\mathcal{H}_0} \right| \\ &\leq K(\mathbf{x}, \mathbf{x})^{\frac{1}{2}} \times \lim_{n \to +\infty} \| f_n \|_{\mathcal{H}_0} \\ &= 0 \,, \end{aligned}$$

showing that $f = 0$. This shows that $\mathcal{H}$ is a pre-Hilbert space with inner product $\langle ., . \rangle_{\mathcal{H}}$. 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

$$\lim_{n \to +\infty} \| f - f_n \|_{\mathcal{H}} = \lim_{n \to +\infty} \lim_{p \to \infty} \| f_p - f_n \|_{\mathcal{H}_0} = 0 \,. \tag{3}$$

This shows in particular that $\mathcal{H}_0$ is dense in $\mathcal{H}$, with respect to the topology defined by the metric $\| . \|_{\mathcal{H}}$. 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 $\| . \|_{\mathcal{H}_0}$ and $\| . \|_{\mathcal{H}}$ coincide on $\mathcal{H}_0$, we then obtain, for any $n, m > N$:

$$\begin{aligned} \| f'_n - f'_m \|_{\mathcal{H}_0} &= \| f'_n - f'_m \|_{\mathcal{H}} \\ &\leq \| f'_n - f_n \|_{\mathcal{H}} + \| f_n - f_m \|_{\mathcal{H}} + \| f_m - f'_m \|_{\mathcal{H}} \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &\leq \varepsilon . \end{aligned} \tag{4}$$

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 $\mathbf{x} \in \mathcal{X}$, by construction, $K_{\mathbf{x}} \in \mathcal{H}_0$ and $\mathcal{H}_0 \subset \mathcal{H}$. To prove (2), let $\mathbf{x} \in \mathcal{X}$ 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

$$\begin{aligned} f(\mathbf{x}) &= \lim_{n\to+\infty} f_n(\mathbf{x}) \\ &= \lim_{n\to+\infty} \langle f_n, K_{\mathbf{x}} \rangle_{\mathcal{H}_0} \\ &= \langle f, K_{\mathbf{x}} \rangle_{\mathcal{H}} . \end{aligned}$$

This concludes the proof of Theorem 1. □

References

[1] N. Aronszajn. Theory of reproducing kernels. *Trans. Am. Math. Soc.*, 68:337 – 404, 1950.