Reproducing Kernel Hilbert Space (RKHS)[1]

The RKHS story is connecting the following objects:

- A non-empty set S;
- A complex-valued function $K = K(t, s)$, $t, s \in S$;
- A Hilbert space H with inner product $(\cdot, \cdot)_H$ and corresponding norm $\|\cdot\|_H$, such that every element of H is a complex-valued function on S.

Definitions

(1) A `kernel` K is a complex-valued, Hermitian, positive semi-definite function on $S \times S$:

$$K(t,s) = \overline{K(s,t)}, \qquad \sum_{k,m=1}^{N} K(t_k, t_m) a_k \overline{a_m} \geq 0,\ t_k \in S,\ a_k \in \mathbb{C}. \tag{1.1}$$

(2) A `reproducing kernel` K on the Hilbert space H is a kernel such that
  - For even $s \in S$, the function $K_t : s \mapsto K(t, s)$ is an element of H;
  - The equality $f(s) = (f, K_s)_H$ holds for every $f \in H$.

(3) The Hilbert space H is called a `reproducing kernel Hilbert space` (RKHS) if there exists a reproducing kernel on H.

Theorem [Aronszajn (1950), Bergman (1950)] The Hilbert space H is a RKHS if and only if, for every $s \in S$, the (linear) functional $f \mapsto f(s)$ is bounded on H: for every $s \in S$, there exists a number $C = C(s)$ such that, for all $f \in H$, $|f(s)| \leq C(s)\|f\|_H$.

Proof. If H is a RKHS with kernel K, then $f(s) = (f, K_s)$, and, by the Cauchy-Schwarz inequality, $|f(s)| \leq \|f\|_H \|K_s\|_H$. Note also that $\|K_s\|_H^2 = \big(K(\cdot, s), K(\cdot, s)\big)_H = \|K(s,s)\|_H^2$. If the point-wise evaluation is a continuous functional, then, by the Riesz representation theorem, $f(s) = (f, g_s)_H$ for some $g_s \in H$, and then $K(t,s) = g_s(t)$.

Corollaries.

(1) For every RKHS H, the corresponding reproducing kernel is unique; if H is separable, with an orthonormal basis $\varphi_m,\ m \geq 1$, then

$$K(t,s) = \sum_{m=1}^{\infty} \varphi_m(t)\overline{\varphi_m(s)}. \tag{1.2}$$

(2) In a RKHS, strong convergence implies point-wise convergence.

(3) A linear subspace of a RKHS is a RKHS.

(4) If a RKHS H is a subspace of a bigger Hilbert space $\widetilde{H}$, then, for $f \in \widetilde{H}$, the function $x \mapsto \big(f, K_x\big)_{\widetilde{H}}$ is the orthogonal projection of f onto H.

(5) If, for a countable collection $t_k \in S$, the collection of functions K_{t_k} is an orthogonal basis in the RKHS H, then, for every $f \in H$,

$$f(s) = \sum_k f(t_k) \frac{K_{t_k}(s)}{K(t_k, t_k)}. \tag{1.3}$$

This observation leads to connections with frames and wavelets.

Examples.

(1) **Moore-Aronszajn theorem:** If K is a kernel (Hermitian positive semi-definite function), then the closure of the linear span of the functions $f(t) = K(t, \cdot)$ with respect to the norm generated by the inner product $\big(f(t), f(s)\big) = K(t,s)$ is a RKHS.

(2) If $\tilde{H}$ is a Hilbert space and A is a bounded operator on $\tilde{H}$ such that $(Ax, y)_{\tilde{H}} = (x, Ay)_{\tilde{H}}$ and $(Ax, x)_{\tilde{H}} \geq 0$, then (taking $S = \tilde{H}$), $K(x,y) = (Ax, y)_{\tilde{H}}$ is a reproducing kernel, and the corresponding RKHS is $\sqrt{A}(\tilde{H})$, where $\sqrt{A}$ is the non-negative symmetric square root of A.

[1]Sergey Lototsky, USC

(3) $L_2((0,1))$ is NOT a RKHS because the functions in $L_2((0,1))$ cannot be evaluated point-wise, and even a continuous function can be changed a lot near one point without changing the L_2 norm that much. The ultimate technicality, though, is that $L_2((0,1))$ is not a collection of function but rather a collection of *equivalence classes*. Constructing a Hilbert space of (bona fide) functions that is not a RKHS is much harder (but possible).

(4) $\dot{H}_1((0,1)) = \{f \in H_1((0,T)) : f(0) = 0\}$ is a RKHS, because $f(t) = \int_0^t f'(r)\,dr$ and so $|f(t)|^2 \le \int_0^1 |f'(r)|^2\,dr$. The corresponding reproducing kernel is $K(t,s) = \min(t,s)$: this is a re-statement of the *original* Cameron-Martin theorem.

(5) The space $A^2(G)$ of functions $f(z) = u(x,y) + \mathfrak{i}v(x,y)$, $\mathfrak{i} = \sqrt{-1}$, analytic in a bounded open sub-set G of the complex plain, with norm $\|f\|^2 = \iint_G |f(z)|^2\,dxdy$ is a RKHS because $\pi r^2|f(z_0)|^2 \le \iint_{|z-z_0|<r} |f(z)|^2\,dxdy$. This is true because the function $|f(z)|^2 = u^2(x,y) + v^2(x,y)$ is sub-harmonic: $\boldsymbol{\Delta}|f(z)|^2 = 2|\nabla u|^2 + 2|\nabla v|^2 \ge 0$, which, in turn is true because the function f is harmonic: $\boldsymbol{\Delta}u = \boldsymbol{\Delta}v = 0$. The corresponding reproducing kernel is called the `Bergman kernel of the domain` G.

(6) The space B_ν of functions f from $L_2(\mathbb{R})$ for which the Fourier transform

$$\hat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\mathfrak{i}yt} f(t)\,dt \tag{1.4}$$

is supported in a fixed interval $[-\nu,\nu]$ is a RKHS; the reproducing kernel is $K(t,s) = (\nu/\pi)\,\mathrm{sinc}(\nu(t-s))$, where $\mathrm{sinc}(x) = \sin(x)/x$. Moreover, taking $t_k \le k\pi/\nu$, $k = 0, \pm 1, \pm 2, \ldots$, equality (1.3) becomes the famous `sampling theorem`

$$f(t) = \sum_k f(t_k)\,\mathrm{sinc}(\nu(t-t_k)).$$

The main point here: by direct computation, the Fourier transform of an indicator function $f_1(t) = \mathbf{1}(-T < t < T)$ is a multiple of the sinc function

$$\hat{f}_1(y) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-\mathfrak{i}yt}\,dt = \sqrt{\frac{2}{\pi}}\,\frac{\sin yT}{y},$$

so that, by duality, the Fourier transform of $\sqrt{2/\pi}\,\nu\,\mathrm{sinc}(\nu t)$ is *exactly* $\mathbf{1}(|y| < \nu)$; as a result, in the frequency domain, the reproducing property

$$f(t) = \int_{\mathbb{R}} f(s)K(t,s)\,ds$$

becomes a trivial identity $\hat{f}(y) = \hat{f}(y)$ and (1.4) becomes the Fourier series expansion of $\hat{f}$.

Historical comments. Originally from Poland, Nachan Aronszajn[2] (1907–1980) had two Ph.D-s.: one from the University of Warsaw and one from the University of Paris under M. Fréchet; after coming to America, he settled at the University of Kansas (1951-77).

Stefan Bergman (1895–1977) was born in the Polish part of the Russian Empire, got his Ph.D. from the University of Berlin under Richard von Mises, later worked in Siberia (Tomsk, 1934-36), Georgia (Tbilisi, 1936-37), and then moved to the US, where he spent some time on the East Coast (MIT, Harvard, Brown, etc.) and the West Coast (Stanford 1952-72); there used to be a special prize in his name from the American Mathematical Society.

While at the University of Chicago, the American mathematician Eliakim Hastings Moore (1862–1932) supervised 31 Ph.D. students, including G. Birkhoff and O. Veblen; he is also the Moore in the Moore-Penrose (generalized) inverse of a matrix.

[2]pronounced close to *aronshine*