Encyclopedia of Mathematics, Supplemental Vol. 3, Kluwer Acad. Publishers, Dordrecht, 2001, 329-331

Reproducing kernel Hilbert Space.

Let H be a Hilbert space of functions defined on an abstract set E.

Let (f, g) denote the inner product and $\| f \| = (f, f)^{\frac{1}{2}}$ is the norm in H. The space H is called RKHS (reproducing kernel Hilbert space) if there exists a function $K(x, y)$ on $E \times E$ such that:

1)$K(x, y) \in H$ for any $y \in E$.

2)$(f(\cdot), K(\cdot, y)) = f(y)$ $\forall f \in H$ (reproducing property).

From this definition it follows that the value $f(y)$ at a point $y \in E$ is a linear continuous functional in H:

$$|f(y)| \leq c(y) \parallel f \parallel, \quad c(y) := \parallel K(\cdot, y) \parallel .$$

The converse is also true. The following theorem holds: *a Hilbert space of functions on a set E is a* RKHS *if and only if* $|f(y)| \leq c(y) \parallel f \parallel$ $\forall y \in E$.

By the Riesz theorem, the above assumption implies the existence of a linear functional $K(\cdot, y)$ such that $f(y) = (f, K(\cdot, y))$. By the construction, the kernel $K(x, y)$ is the reproducing kernel for H.

An example of a construction of a RKHS is the rigged triple of Hilbert spaces $H_+ \subset H_0 \subset H_-$, which is defined as follows [5]. Let H_0 be a Hilbert space of functions, $A > 0$ be a linear densely defined selfadjoint operator on H_0, $A\varphi_j = \lambda_j \varphi_j$, (the eigenvalues $\lambda_j > 0$ are counted according to their multiplicities and $||\varphi_j||_0 = 1$), and assume that

$$\Lambda^2 := \sum_{j=1}^{\infty} \lambda_j < \infty, \quad |\varphi_j(x)| < c \quad \forall j, x.$$

Define $H_- \supset H_0$, a Hilbert space with the inner product $(u, v)_- := (A^{\frac{1}{2}} u, A^{\frac{1}{2}} v)_0$, H_- is the completion of H_0 in the norm $\parallel u \parallel := (u, u)_-^{\frac{1}{2}}$. Let $H_+ \subset H_0$ be the dual space to H_- with respect to H_0. Then the inner product in H_+ is defined by the formula $(u, v)_+ := (A^{-\frac{1}{2}} u, A^{-\frac{1}{2}} v)_0$, and $H_+ = R(A^{\frac{1}{2}})$, equipped with the inner product $(u, v)_+$, is a Hilbert space.

Define $B(x, y) = \sum_{j=1}^{\infty} \lambda_j \varphi_j(x) \overline{\varphi_j(y)}$, where the overline stands for complex conjugate. For any y, one has $B(x, y) \in H_+$. Indeed

$$\parallel B(x, y) \parallel_+ = \parallel \sum_{j=1}^{\infty} \lambda_j^{\frac{1}{2}} \varphi_j(x) \overline{\varphi_j(y)} \parallel_0 \leq c \sum_{j=1}^{\infty} \lambda_j = c\Lambda^2 < \infty.$$

Furthermore

$$(u, B(x, y))_+ = (u, A^{-1}B) = u(y),$$

so that $B(x, y)$ is the <u>reproducing</u> <u>kernel</u> in H_+. Moreover $|u(y)| \leq c(y) \parallel u \parallel_+$, where $c(y) > 0$ is a constant independent of $u \in H_+$. Indeed, if $u_j := (u, \varphi_j)_0$ and $u \in H_+$, then $u = A^{\frac{1}{2}}v$, $v \in H_0$ $v_j\lambda_j^{\frac{1}{2}} = u_j$, and $|u(y)| \leq \sum_{j=1}^{\infty} |u_j\varphi_j(y)| \leq c\Lambda \parallel v \parallel = c\Lambda \parallel u \parallel_+$.

Thus H_+ is a RKHS with the reproducing kernel $B(x, y)$ defined above. If $K(x, y)$ is a function on $E \times E$, such that

$$\sum_{i,j=1}^{n} K(x_i, x_j)t_j\overline{t_i} \geq 0 \quad \forall t \in \mathbb{C}^n, \quad \forall x_i \in E, \tag{1}$$

then one can define a pre-Hilbert space H^0 of functions of the form

$$f(x) := \sum_{j=1}^{J} K(x, y_j)c_j, \quad c_j = \text{ const}\,.$$

The inner product of the two functions from H^0 is defined by the formula

$$(f, g) := \left(\sum_{j=1}^{J} K(x, y_j)c_j, \quad \sum_{m=1}^{M} K(x, z_m)\beta_m\right) = \sum_{j,m} K(z_m, y_j)c_j\overline{\beta_m}.$$

This definition makes sense because of (1) and because of reproducing property 2). In particular, $(f, f) \geq 0$, as follows from (1), and if $(f, f) = 0$ then $f = 0$, as follows from property 2).

Indeed,

$$(f(x), K(x, y)) = \left(\sum_{j=1}^{J} K(x, y_j)c_j, K(x, y)\right) = \sum_{j=1}^{J} K(y, y_j)c_j = f(y) \quad \forall y \in E.$$

Thus, if $(f, f) = 0$, then $\parallel f \parallel = 0$ and $|f(y)| \leq \parallel f \parallel \parallel K(x, y) \parallel = 0$, so $f(y) = 0$ as claimed.

Denote by H the completion of H^0 in the norm $||f||$. Then H is a reproducing kernel Hilbert space (RKHS) and $K(x, y)$ is its reproducing kernel (rk).

The RKHS is uniquely defined by its rk. Indeed, if H_1 is another RKHS with the same rk $K(x, y)$, then $H^0 \subset H_1$ and H^0 is dense in H_1: if $f \in H_1$, $0 = (f, K(x, y))_1 = f(y)$ $\forall y \in E$, then $f \equiv 0$. Using this and the equality $(f, g)_{H_1} = (f, g)_H$ for all $f, g \in H^0$, one can check that $H \subset H_1$ and vice versa, so $H = H_1$, that is, H and H_1 consist of the same set of elements. Moreover, the norms in H and H_1 are equal. Let us check this. Take an arbitrary $f \in H_1$ and a sequence $f_n \in H^0$, $\parallel f_n - f \parallel_1 \to 0$. Then

$$\parallel f \parallel_1^2 = \lim_{n\to\infty} \parallel f_n \parallel_1^2 = \lim_{n\to 0} \left(\sum_{j_n=1}^{J_n} K(x, y_{j_n})c_{j_n}, \sum_{m_n=1}^{J_n} K(x, y_{m_n})c_{m_n}\right)_1 =$$

$$\sum_{j_n,m_n}^{J_n} K(y_{m_n}, y_{j_n}) c_{j_n} \overline{c_{m_n}} = \lim_{n\to\infty} (f_n, f_n) = ||f||^2.$$

Thus the norms in H_1 and H are equal, as claimed, and so are the inner products (by the polarization identity).

Define a linear operator $L : \mathcal{H} \to H$, $D(L) = \mathcal{H}$, where $\mathcal{H} = L^2(T, dm)$ and H is the range $R(L)$ of L, which we equip with the Hilbert space structure later:

$$f(x) = LF := \int_T F(t)\overline{h(t,x)}dm(t). \tag{2}$$

Here T is a domain in $\mathbf{R^n}$ and m is a positive measure on T, $m(T) < \infty$, $h(t,x) \in \mathcal{H}$ $\forall x \in E$, and we assume that L is *injective,* that is, the system $\{h(t,x)\}_{\forall x \in E}$ is total in $\mathcal{H}$.

Define

$$K(x,y) := \int_T h(t,y)\overline{h(t,x)}dm(t) = (h(\cdot,y), h(\cdot,x))_{\mathcal{H}} \tag{3}$$

This kernel clearly satisfies condition (1) and therefore is a rk for the RKHS H_K which it generates. Clearly $K(x,y) \in H$ $\forall y \in E$. If $f \in H$, that is, $f = LF$, $f \in \mathcal{H}$, then

$$(f(\cdot), K(\cdot,y))_H = (LF, K(\cdot,y))_H = ((F(\cdot), h(\cdot,x))_{\mathcal{H}}, (h(\cdot\cdot,y), h(\cdot\cdot,x))_{\mathcal{H}})_H =$$

$$(F(\cdot), (h(\cdot\cdot,y), (h(\cdot,x), h(\cdot\cdot,x))_H)_{\mathcal{H}})_{\mathcal{H}} = (F(\cdot), h(\cdot,y))_{\mathcal{H}} = f(y),$$

if one equips H with the inner product such that $(f,g)_H = (F,G)_{\mathcal{H}}$. This requirement is formally equivalent to the following one: $(h(s,x), h(t,x))_H = \delta_m(t-s)$, where $(h(s,y), \delta_m(t-s))_{\mathcal{H}} = h(t,y)$, so that the distributional kernel $\delta_m(t-s)$ is not the usual delta-function, but the one which acts by the rule $\int_T dm(t)F(t)\int_T dm(s)G(s)\delta_m(t-s) = \int_T dm(t)F(t)G(t)$, and formally one has $\int_T dm(s)G(s)\delta_m(t-s) = G(t)$.

With the inner product $(f,g)_H$ the linear set $R(L)$ becomes a Hilbert space:

$$\begin{aligned}(f,g)_H = (LF, LG)_H &= \int_T\int_T dm(t)dm(s)F(t)\overline{G(s)}(h(s,x), h(t,x))_H \\ &= \int_T dm(t)F(t)\overline{G(t)} = (F,G)_{\mathcal{H}}.\end{aligned} \tag{4}$$

Thus, this inner product makes L an isometry defined on all of $\mathcal{H}$ and makes $H = R(L)$ a (complete) Hilbert space, namely $H = H_K$, a RKHS. Since L is assumed injective, it follows that L^{-1} is defined on all of $R(L) = H$ and, since H is complete in the norm $||f|| = (f,f)_H^{\frac{1}{2}}$, one concludes that L^{-1} is continuous (by the Banach theorem). Consequently, L is a coisometry, that is $L^* = L^{-1}$, where L^* is the adjoint operator to L. If $L^* = L^{-1}$, then one can write an inversion formula for the linear transform L similar to the well-known inversion formula for the Fourier transform. Formally one has:

$$f(x) = (F(t), h(t,x))_{\mathcal{H}}, \quad (f(x), h(s,x))_H = F(s).$$

The space $H = H_K$ is the RKHS generated by kernel (3) which is the rk for H. The above formal inversion formulas may be of practical interest if the norm in H is a standard one. In this case the second formula should be suitably interpreted since $F(s)$ is defined at m-almost all s.

In [6] it is claimed that the characterization of the range of the linear operator L, defined in (3), can be given as follows: $R(L) = H_K$, where H_K is the RKHS generated by kernel (3).

However, in fact such a characterization does not give, in general, a practically useful necessary and sufficient conditions for $f(x) \in R(L)$ because the norm in H_K is not defined in terms of the standard norms such as Sobolev or Hölder ones (see [3]- [5]). However, in the cases when the norm in H_K is equivalent to a standard norm, the above characterization becomes efficient (see [3]- [5], and also [6]).

Many concrete examples of the RKHS can be found in [1],[2] and [6].

Papers [1] and [7] are important in this area, book [6] contains many references, and book [2] is an earlier book important for the development of the theory of RKHS.

References

[1] Aronszajn, N., Theory of reproducing kernels, *Trans. Amer. Math. Soc.*, 68, (1950), 337-404.

[2] Bergman, S., The kernel function and conformal mapping, *Amer. Math. Soc.*, Providence, RI, 1950.

[3] Ramm, A.G., On the theory of reproducing kernel Hilbert spaces, *Jour. of Inverse and Ill-Posed Probl.*, 6, N5 (1998), 515-520.

[4] Ramm, A.G., On Saitoh's characterization of the range of linear transforms, In: Inverse Problems, Tomography and Image Processing, Plenum, New York, 1998, 125-128. (ed. A.G. Ramm)

[5] Ramm, A.G., Random fields estimation theory, Longman/Wiley, New York, 1990.

[6] Saitoh, S., Integral transforms, reproducing kernels and their applications, *Pitman Res. Notes*, Longman, New York, 1997.

[7] Schwartz, L., Sous-espaces hilbertiens d'espaces vectoriels topologiques et noyaux associès (noyaux reproduisants), *J. Analyse Math.*, 13, (1964), 115-256.

A.G. Ramm
Mathematics Department
Kansas State University,
Manhattan, KS 66506-2602, USA
ramm@math.ksu.edu