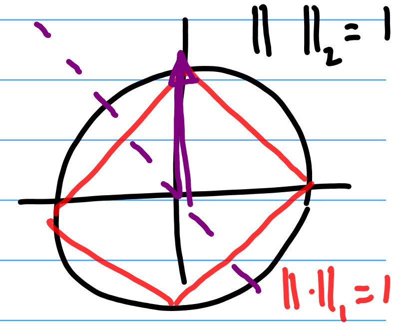
Lecture 5: From Riesz Representation to RKHSs

Question from last lecture, regarding Rep. thm.

Given
$$\min_{h \in H} L\Big(\sum_{j=1}^{n} \langle h, h_j \rangle\Big) + R(\|h\|)$$
why is it that we cannot replace $\|\cdot\|$ with an equivalent norm? Recall, from the proof we used Pythagorean ident. to show that $\|h\| \geq \|h''\|$. This is not true for equivalent norms eg.



5.1 Proof of Riesz's Rep. Thm

Recall Riesz's Rep. Thm.

Let H be a Hilbert space & let $\phi \in H^*$

Then $\phi(h) = \langle h, \hat{\phi} \rangle$ where $\hat{\phi} \in H$ is uniquely determined by ϕ & $\|\hat{\phi}\| = \|\phi\|_*$

<u>Proof</u>:- The proof follows three steps

(i) show $\phi(h) = \langle \hat{\phi}, h \rangle$ holds

(ii) Show $\hat{\phi}$ is unique

(iii) show $\|\hat{\phi}\| = \|\phi\|_*$

(ii) If $f = 0$ then we simply take $\phi = 0$ & we are done. So let $f \neq 0$.

We need some basic results from the theory of bdd lin. operators on Hilbert spaces:

(Scribe, add Cor 2.7-10 & Thm 3.3-4)

$\text{Null}(\phi)$ is a closed subspace of H.

Now, since $\phi \neq 0$ then $\text{Null}(\phi) \neq H$ & so $\text{Null}(\phi)^{\perp} \neq \{0\}$. In otherwords we can pick an element $h_0 \in \text{Null}(\phi)^{\perp}$ st. $\phi(h_0) \neq 0$.

Now pick an arbitrary element $h \in H$ & set

$$v = \phi(h) h_0 - \phi(h_0) h$$

Observe that $\phi(v) = \phi(h)\phi(h_0) - \phi(h_0)\phi(h) = 0$ so that $v \in \text{Null}(\phi)$!

Since we assume $h_0 \in \text{Null}(\phi)^{\perp}$ then $\langle h, v \rangle = 0$!

$$0 = \langle h_0, v \rangle = \phi(x) \|h_0\|^2 - \phi(h_0) \langle h_0, h \rangle$$

Solve for $\phi(h)$ to get

$$\phi(h) = \frac{\phi(h_0) \langle h_0, h \rangle}{\|h_0\|^2}$$

$$= \langle \hat{\phi}, h \rangle \text{ when } \hat{\phi} = \frac{\phi(h_0)}{\|h_0\|^2} h_0$$

Since h was arbitrary we are done.

(ii) Suppose $\hat{\phi}$ & $\hat{\phi}'$ are two representes of ϕ then $\phi(h) = \langle \hat{\phi}, h \rangle = \langle \hat{\phi}', h \rangle$

Then $\langle \hat{\phi} - \hat{\phi}', h \rangle = 0 \quad \forall h \in H$. Let $h = \hat{\phi} - \tilde{\phi}'$

Then $\|\hat{\phi} - \hat{\phi}'\|^2 = 0 \implies \hat{\phi} = \hat{\phi}'$ Contradiction!

(iii) If $\phi = 0$ we take $\hat{\phi} = 0$ & $\|\phi\|_* = \|\hat{\phi}\| = 0$.

So suppose $\phi \neq 0$. Since $\phi \in H^*$ then

$$\|\hat{\phi}\|^2 = \langle \hat{\phi}, \hat{\phi} \rangle = \phi(\hat{\phi}) \leq \|\phi\|_* \cdot \|\hat{\phi}\|$$

$$\implies \|\hat{\phi}\| \leq \|\phi\|_*.$$

Conversely, $\forall h \in H$ s.t. $\|h\| = 1$ we have

$$|\phi(h)| = |\langle \hat{\phi}, h \rangle| \leq \|\hat{\phi}\|$$

so that $\|\phi\|_* \leq \|\hat{\phi}\|$.

Thus, $\|\hat{\phi}\| = \|\phi\|_*$

■

52 First Introduction of RKHSs

Let $\Omega \subseteq \mathbb{R}^d$ & let H be a Hilbert space of functions on Ω; $f \in H$, $f : \Omega \to \mathbb{R}$.

Moreover, suppose that the pointwise eval. functional $\delta_x \in H^*$, ie,

$$\delta_x(f) = f(x) \qquad \forall x \in \Omega$$

$$|\delta_x(f)| \leq c \|f\|$$

Then by Riesz's Rep. thm. we have that

$$f(x) = \delta_x(f) = \langle K_x, f \rangle$$

where $K_x \in H$ is the rep. of δ_x. At the same time since $K_x \in H$ we have.

$$K_x(y) = \delta_y(K_x) = \langle K_y, K_x \rangle$$

$$= \langle K_x, K_y \rangle = \delta_x(K_y) = K_y(x)$$

In otherwords we can define

$$K(x,y) = K_x(y) \qquad \forall x,y \in \Omega$$

The function K is called a Kernel.

We already verified that K is symmetric

$$K(x,y) = K(y,x)$$

At the same time, we have for any $f \in H$

$$\langle K(x,\cdot), f \rangle = \langle k_x, f \rangle = f(x)$$

This is called the Reproducing property of the kernel K.

<u>Def^n</u> A Hilbert space H of functions from a set Ω to $\mathbb{R}$ is called a Reproducing kernel Hilbert space (RKHS) if Pointwise evaluation is a bdd lin. func. on H.

We already showed that K is symm. & satisfies the reproducing prop. We can indeed show a little more

Since $\mathcal{H}$ is a vector space take

$$H \ni f = \sum_{j=1}^{n} \xi_j K(x_j, \cdot) \quad \text{for } \{x_1, \dots, x_n\} \subset \Omega$$
$$\{\xi_1, \dots, \xi_n\} \subset \mathbb{R}$$

Then we can compute

$$0 \leq \|f\| = \langle f, f \rangle = \left\langle \sum_{j=1}^{n} \xi_j K(x_j, \cdot), \sum_{k=1}^{n} \xi_k K(x_k, \cdot) \right\rangle$$
$$= \sum_{j=1}^{n} \sum_{k=1}^{n} \xi_j \xi_k \langle K(x_j, \cdot), K(x_k, \cdot) \rangle$$
$$= \sum_{j=1}^{n} \sum_{k=1}^{n} \xi_j \xi_k K(x_j, x_k)$$

Thus, we infer that K is positive definite in the following sense.

Def^n A kernel $K : \Omega \times \Omega \to \mathbb{R}$ is positive definite if for all $n \in \mathbb{N}$, $\underline{\xi} \in \mathbb{R}^n$ & collection $X = \{x_1, \dots, x_n\} \subset \Omega$ it holds that

$$\underline{\xi}^T K(X, X) \underline{\xi} \geq 0$$

where we introduced the shorthand notation $(K(X,X))_{ij} = K(x_i, x_j)$

The matrix $K(X,X)$ is often referred to as the kernel matrix.

In summary we showed, using Riesz's rep:

- Any Hilbert space of functions where pointwise eval. is a bdd. lin. func. is an RKHS.

- The rep. k_x of δ_x defines a kernel $K: \Omega \times \Omega \to \mathbb{R}$ that satisfies the reproducing property

$$f(x) = \langle f, K(x, \cdot) \rangle$$

- The kernel K is positive definite & symmetric (PDS).