Lecture 22
Riesz Representation

John McCuan

March 28, 2020

We are now going to prove the Riesz representation theorem for bounded linear functionals on a Hilbert space. Let me recall the statement:

Theorem 1 *(Riesz representation theorem) If* $\langle \, \cdot \, , \, \cdot \, \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ *is* **any** *inner product on a Hilbert space* $\mathcal{H}$ *and* $L : \mathcal{H} \to \mathbb{R}$ *is a bounded linear functional, i.e.,* $L \in \beth^0(\mathcal{H}) = \mathcal{H}^*$*, then there exists a unique* $u \in \mathcal{H}$ *such that*

$$Lv = \langle u, v \rangle \qquad \text{for all } v \in \mathcal{H}.$$

The last lecture should have given you a pretty good idea about how to prove this, or at least start. You start by looking at the null space $\ker(L)$ of L which is a closed subspace of $\mathcal{H}$.

Exercise 1 *Show the null space* $\ker(L)$ *of a bounded linear functional* $L \in \mathcal{H}^*$ *is closed.*

We even had a prospective choice for the representing vector which we got like this: Take $u_0 \in \ker(L)^\perp \backslash \{\mathbf{0}\}$. Then (according to Riesz) some scaling $u = \alpha u_0$ will work. If so, then

$$\alpha = \frac{Lu_0}{|u_0|^2}. \tag{1}$$

But (hopefully) the big question which should have been left in your mind at the end of the last lecture was the following:

If we look at

$$Lv = L\left(v - \frac{\langle u_0, v\rangle}{|u_0|^2}\,u_0\right) + L\left(\frac{\langle u_0, v\rangle}{|u_0|^2}\,u_0\right)$$

while I certainly know $v - \langle u_0, v\rangle\, u_0/|u_0|^2 \in \text{span}\{u_0\}^\perp$, how do I know

$$L\left(v - \frac{\langle u_0, v\rangle}{|u_0|^2}\,u_0\right) = Lv - \frac{\langle u_0, v\rangle}{|u_0|^2}\,Lu_0 = Lv - \left\langle \frac{Lu_0}{|u_0|^2}\,u_0, v\right\rangle = 0? \quad (2)$$

If we have the equality in (2), then we can say

$$Lv = L\left(\frac{\langle u_0, v\rangle}{|u_0|^2}\,u_0\right) = \frac{\langle u_0, v\rangle}{|u_0|^2}\,Lu_0 = \left\langle \frac{Lu_0}{|u_0|^2}\,u_0, v\right\rangle = \langle \alpha u_0, v\rangle \tag{3}$$

so we have representation.

Let me point out that we can certainly take $u_0 \in \ker(L)^\perp\backslash\{\mathbf{0}\}$. At least if we can't, then $\ker(L) = \mathcal{H}$ and we can take $u = \mathbf{0} \in \mathcal{H}$, and we are done.

Exercise 2 *Once we have a vector $u_0 \in \ker(L)\backslash\{\mathbf{0}\}$, then we can assume $|u_0| = 1$, so $\alpha = Lu_0$. How does this simplify the main question (2) above?*

Still the question of (2) remains.

You might also recall that the equality of (2) would follow if we knew $\text{span}\{u_0\}^\perp \subset \ker(L)^{\perp\perp} = \ker(L)$. The set equality $V^{\perp\perp} = V$ is true for any closed subspace in a Hilbert space, but we do not have a proof of that fact. (The analogue in the finite dimensional case used a **direct sum decomposition**, and the proof of the direct sum decomposition lemma used that we had a finite basis.) We are apparently stuck.

Riesz' Trick in $\mathbb{R}^n$

Let us return briefly to our consideration of the finite dimensional case $\ell : \mathbb{R}^n \to \mathbb{R}$ and even to our specific mapping $\ell : \mathbb{R}^3 \to \mathbb{R}$ determined by $\mathbf{e}_j \mapsto j-1$ for $j = 1, 2, 3$. The argument from the last lecture should give us that the vector

$$\mathbf{z} = \mathbf{v} - \frac{\mathbf{v}\cdot\mathbf{u}_0}{|\mathbf{u}_0|}\,\frac{\mathbf{u}_0}{|\mathbf{u}_0|} \tag{4}$$

from the Gram-Schmidt procedure satisfies $\mathbf{z} \in \ker(\ell)$. It will be observed, however, that this is hardly the case. In fact, I see no clear and obvious way to see

$$\ell(\mathbf{z}) = \ell(\mathbf{v}) - \frac{\mathbf{v}\cdot\mathbf{u}_0}{|\mathbf{u}_0|}\,\frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|} = 0.$$

Of course, you can write $\mathbf{z}$ as $\mathbf{z} = \mathbf{z}_0 + \mathbf{w}_0$ with $\mathbf{z}_0 \in \ker(\ell)$ and $\mathbf{w}_0 \in \ker(\ell)^\perp$ and then follow through the argument of the double orthgonal complement lemma. Even if you do that to conclude $\mathbf{z} = \mathbf{z}_0 \in \ker(\ell)$, the relation between the decomposition in (4) involving $\mathbf{v}$ and $\mathbf{u}_0$ and the decomposition $\mathbf{z} = \mathbf{z}_0 + \mathbf{w}_0$ remains (as far as I can see) still rather obscure. In particular, the relationship depends on the use of some basis for $V = \ker(\ell)$ which you don't even have (at least easily) in the infinite dimensional case.

Riesz had the very clever (and elegant) idea of decomposing $\mathbf{v}$ in a manner similar to

$$\mathbf{v} = \mathbf{v} - \frac{\mathbf{v}\cdot\mathbf{u}_0}{|\mathbf{u}|^2}\,\mathbf{u}_0 \qquad \text{with} \qquad \mathbf{z} = \frac{\mathbf{v}\cdot\mathbf{u}_0}{|\mathbf{u}|^2}\,\mathbf{u}_0$$

but in a crucially different way using a *different multiple of* $\mathbf{u}_0$. That is, consider

$$\mathbf{v} = (\mathbf{v} - \beta\mathbf{u}_0) + \beta\mathbf{u}_0 \tag{5}$$

where β is some constant to be determined. This gives the residual vector

$$\mathbf{z} = \mathbf{v} - \beta\mathbf{u}_0 \qquad \text{with} \qquad \ell(\mathbf{z}) = \ell(\mathbf{v}) - \beta\ell(\mathbf{u}_0). \tag{6}$$

Now, all we need to know is that $\mathbf{u}_0 \notin \ker(\ell)$ and we can take $\beta = \ell(\mathbf{v})/\ell(\mathbf{u}_0)$. Then we get $\ell(\mathbf{z}) = 0$, i.e., $\mathbf{z} \in \ker(\ell)$, automatically. In this way we have $\ell(\mathbf{v}) = \beta\ell(\mathbf{u}_0)$. But, do we get Riesz representation from this and the choice suggested in (1)? In fact,

$$\mathbf{v}\cdot\frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\,\mathbf{u}_0 = (\mathbf{v} - \beta\mathbf{u}_0)\cdot\frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\,\mathbf{u}_0 + \beta\mathbf{u}_0\cdot\frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\,\mathbf{u}_0 = \beta\,\ell(\mathbf{u}_0) = \ell(\mathbf{v})$$

since $\mathbf{v} - \beta\mathbf{u}_0 \in \ker(\ell)$ and $\mathbf{u}_0 \in \ker(\ell)^\perp$.

Proof of the Riesz Representation Theorem

Consider

$$\mathcal{N} = \{z \in \mathcal{H} : Lz = 0\}$$

the **null space** of L. This is a closed vector subspace of $\mathcal{H}$.

If $\mathcal{N} = \mathcal{H}$, then we have representation using the zero vector:

$$Lv = \langle \mathbf{0}, v\rangle,$$

and the representation is unique since

$$\langle u, v\rangle = 0 \quad \text{for all } v \in V \qquad \Longrightarrow \qquad \|u\|^2 = \langle u, u\rangle = 0.$$

So, obviously the more interesting case is when $\mathcal{N} \subsetneq \mathcal{H}$. In this case, we can take a vector $u_0 \in \mathcal{N}^\perp \backslash \{0\}$. Now, normally one would think there are many such vectors u_0, so it's not so obvious that you can count on anything special from this one. Apparently, however, the fact that the image of the functional L is $\mathbb{R}$ (which is conspicuously one dimensional) somehow narrows the possibilities.

Thus, the crucial ansatz is to look for a scaling of u_0 as the choice of u. That is, we look for some $\alpha \in \mathbb{R}$ for which

$$Lv = \langle \alpha u_0, v\rangle \qquad \text{for all } v \in \mathcal{H}.$$

Once this ansatz is written down, then you know the identity of the scalar α because you must have

$$Lu_0 = \alpha \|u_0\|^2.$$

That is,

$$\alpha = \frac{Lu_0}{\|u_0\|^2}.$$

Once that determination is made, one simply needs to see if (or show that)

$$u = \frac{Lu_0}{\|u_0\|^2}\, u_0 \qquad \text{works.}$$

The idea for accomplishing this is somewhat reminiscent of the decomposition of a vector in a direct sum. That is, we take an arbitrary vector v and decompose it in terms of a component along $\mathcal{N}^\perp$ or more precisely along $\text{span}(u_0)$:

$$v = (v - \beta u_0) + \beta u_0.$$

In order to have $v - \beta u_0 \in \mathcal{N}$, we take $\beta = Lv/Lu_0$. This is a well-defined vector since $u_0 \notin \mathcal{N}$ and so $Lu_0 \neq 0$. Note then, that

$$L(v - \beta u_0) = 0, \qquad \text{so} \qquad v - \beta u_0 \in \mathcal{N}, \qquad \text{and} \qquad \langle u_0, v - \beta u_0\rangle = 0.$$

Then we can compute

$$\begin{aligned}\langle \alpha u_0, v\rangle &= \langle \alpha u_0, v - \beta u_0\rangle + \langle \alpha u_0, \beta u_0\rangle \\ &= \langle \alpha u_0, \beta u_0\rangle \\ &= \alpha\beta \|u_0\|^2 \\ &= Lv\end{aligned}$$

since

$$\alpha\beta = \frac{Lu_0}{\|u_0\|^2}\frac{Lv}{Lu_0} = \frac{Lv}{\|u_0\|^2}.$$

Thus, we have existence of a vector $u = \alpha u_0$ for which

$$\langle u, v \rangle = Lv \qquad \text{for all } v \in V.$$

Uniqneness is, again rather easy: If $\langle u, v\rangle = \langle \tilde{u}, v\rangle$, then taking $v = u - \tilde{u}$, we get

$$\|u - \tilde{u}\|^2 = \langle u - \tilde{u}, u - \tilde{u}\rangle = 0. \qquad \square$$

One thing to note about our discussion of the Riesz representation theorem: Essentially no inequalities, estimates, questions of convergence, or limits were used. Basically no (hard) analysis was mentioned. There was a good deal of (perhaps tricky) algebra and especially linear algebra. To be fair, the usual proofs that a **closest vector** in a subspace to a given vector outside that subspace exists involve showing some sequence is Cauchy, and that involves some elementary estimation, but that is about it. Incidentally, this is the point where the completeness of a Hilbert space comes in.

Exercise 3 *Show that a closed subset of a Hilbert space is metrically complete. In particular a closed subspace of a Hilbert space is a Hilbert space. Give an example of a subspace of a Hilbert space which is neither closed nor complete.*

Exercise 4 *Give a simplified version of the proof of the Riesz represenation theorem utilizing the remark made above that you can assume $|u_0| = 1$.*

Summary: Final Remarks

Once we have proved the Riesz representation theorem in Hilbert space, we can look back and see

$$\ker(\ell) = \{\mathbf{z} \in \mathcal{H} : \ell(\mathbf{z}) = 0\} = \{\mathbf{z} \in \mathcal{H} : \langle \mathbf{u}, \mathbf{z}\rangle = 0\} = \text{span}\{\mathbf{u}\}^\perp.$$

Thus, since the span $\text{span}\{\mathbf{u}\}$ of the representing vector is one-dimensional, the space $\ker(\ell)$ is the (very large) orthogonal complement. Our intuition was correct, even though we didn't use it directly in the proof.

It will noted that all estimates involved here (all analysis) has been swept under the rug in the exercise from the previous lecture in which one shows the projection onto a closed subspace is well-defined. That analysis, moreover, is aimed at showing the existence of a vector closest to the closed linear subspace which is the null space of the function ℓ. In our application to existence and uniqueness of weak solutions of Poisson's equation, the actual operator from the PDE is actually further hidden away in the inner product, and we only really see the abstract properties of the inner product in the proof of the Riesz theorem above. In this way, there is a kind of double sweeping of the analysis under the rug. The only means to get back to what is actually happening with the operator, in this case the Laplace operator, is in regard to the Poincaré inequality which essentially renders the weak adjoint

$$B[u,v] = \int Du \cdot Dv \qquad \text{with} \qquad [u]_{W^{1,1}} = B[u,u]$$

an inner product.

Exercise 5 (challenge) *Where is the positive definiteness of the inner product used/required in the proof of* **existence** *in the Riesz representation theorem?*