Lecture 21
Riesz Representation on $\mathbb{R}^n$

John McCuan

March 28, 2020

We will prove the Riesz representation theorem for bounded linear functionals on a Hilbert space below,[1] and this is the setting where the result is of primary interest. In particular, we want to apply the result to get the existence and uniqueness of weak solutions of Poisson's PDE with homogeneous boundary conditions, and in that case our Hilbert space will be $\mathcal{H} = H_0^1(U)$ which is a Sobolev space. Here is 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*

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

It may be helpful to consider the special case of this result in the familiar setting of finite dimensional linear mappings. Say we have a linear function

$$\ell : \mathbb{R}^n \to \mathbb{R},$$

that is, a **linear functional**. Such a function is always continuous and has a number of other properties with which are (or should be) quite familiar from linear algebra. One of these is that there exists a particular vector $\mathbf{u} \in \mathbb{R}^n$ such that

$$\ell(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \qquad \text{for all } \mathbf{v} \in \mathbb{R}^n. \tag{1}$$

This observation, and the fact that the vector $\mathbf{u}$ is uniquely determined by the functional ℓ, constitute the essential assertion of the Riesz representation theorem. The

[1] More properly in the next lecture.

linear functional ℓ is said to be "represented" by the vector $\mathbf{u}$. Of course, the linear functional ℓ is also represented in these theorems **by the inner product**, and that (though it usually almost goes without saying and the emphasis is put on the representing vector $\mathbf{u}$) is actually a really important point for us in our application.

The right side of the relation (1) uses the usual Euclidean dot product, and the Riesz representation theorem essentially replaces this with an arbitrary inner product (and applies to bounded linear functionals on a (potentially) infinite dimensional Hilbert space).

Let's attempt to think a little bit about the vector $\mathbf{u}$ and why this kind of representation occurs. The usual way to do this is to note (or observe) that

$$\mathbf{u} = (\ell(\mathbf{e}_1), \ell(\mathbf{e}_2), \dots, \ell(\mathbf{e}_n))$$

where $\mathbf{e}_j$ is the ***j*-th standard unit basis vector** (with 1 in the j-th entry and zeros in all other entries). This choice clearly gives

$$\ell(\mathbf{e}_j) = \mathbf{e}_j \cdot \mathbf{u} \qquad \text{for } j = 1, 2, \dots, n.$$

The identity

$$\ell(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \qquad \text{for every } \mathbf{v} \in \mathbb{R}^n$$

follows by linearity since we can write $\mathbf{v} = \sum (\mathbf{v} \cdot \mathbf{e}_j)\, \mathbf{e}_j$.

In an infinite dimensional inner product space, like $L^2(a, b)$, if you consider an orthonormal basis like $\{\sin j\pi(x-a)/(b-a)\}_{j=1}^{\infty}$, then you will have to consider series representation, which can be done, but then you will have questions of convergence and other inconveniences to deal with. In some sense, it is the point of **functional analysis** to avoid such messy details and give a fundamentally different argument avoiding reference to a basis, or limits, or estimates—but sticking to ideas of linear algebra. Let's see if we can suggest how this might be done in the finite dimensional case where we understand everything.

First of all, if we have a representation $\ell(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$, then while the connection between ℓ and $\mathbf{u}$ may not be at all obvious, there is one connection which is easy to make: The subspace

$$\ker(\ell) = \{\mathbf{x} : \ell(\mathbf{x}) = 0\}$$

which is the **kernel** or **null space** of ℓ will consist of the vectors which are perpendicular to $\mathbf{u}$. That is, on the one hand $\ker(\ell)$ is a vector subspace of $\mathbb{R}^n$ and, on the other hand, the **orthogonal complement** of any vector $\mathbf{u}$

$$\mathbf{u}^{\perp} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{u} = 0\}$$

is also a subspace. If our representation construction is going to work, **these subspaces must match**. Note that this gives us a place to start without ever mentioning a basis.

If $\ker(\ell) = \mathbb{R}^n$, then we can get a representation using $\mathbf{u} = \mathbf{0}$ (the zero vector). In fact, in this case, $\mathbf{u} = \mathbf{0}$ is the unique choice since

$$\mathbf{u} \cdot \mathbf{v} = \tilde{\mathbf{u}} \cdot \mathbf{v} \quad \text{for all } \mathbf{v} \qquad \Longrightarrow \qquad (\mathbf{u} - \tilde{\mathbf{u}}) \cdot \mathbf{v} = 0.$$

And taking $\mathbf{v} = \mathbf{u} - \tilde{\mathbf{u}}$, we get $|\mathbf{u} - \tilde{\mathbf{u}}| = 0$.

Otherwise, there is some **nonzero** vector $\mathbf{u}_0 \in \ker(\ell)^\perp$. The basic observation of the Riesz theorem is then that **some scaling $\mathbf{u} = \alpha \mathbf{u}_0$ of $\mathbf{u}_0$ will work**. If that is correct, then we need

$$\ell(\alpha \mathbf{u}_0) = \alpha \mathbf{u}_0 \cdot (\alpha \mathbf{u}_0) = \alpha^2 |\mathbf{u}_0|^2,$$

so $\alpha = \ell(\mathbf{u}_0)/|\mathbf{u}_0|^2$. Then the question is: Will

$$\mathbf{u} = \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\, \mathbf{u}_0 = \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|}\, \frac{\mathbf{u}_0}{|\mathbf{u}_0|} \tag{2}$$

work?

In fact, we will show that this does work, even in the general infinite dimensional case. Before we do that, however, it may be instructive to consider a very specific example. Let $\ell : \mathbb{R}^3 \to \mathbb{R}^1$ by determined by

$$\begin{cases} \mathbf{e}_1 \mapsto 0 \\ \mathbf{e}_2 \mapsto 1 \\ \mathbf{e}_3 \mapsto 2. \end{cases} \tag{3}$$

The linearity argument above tells us $\mathbf{u} = (0, 1, 2)$ should be the unique vector such that $\ell(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$, and clearly this works.

We are supposed, however, to see/find this vector $\mathbf{u} = (0, 1, 2)$ without using the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. To do so, we look at

$$\ker(\ell) = \{\mathbf{z} \in \mathbb{R}^3 : \ell(\mathbf{z}) = 0\}.$$

We might be inclined, looking at the definition of ℓ in (3), to think

$$\ker(\ell) = \text{span}\{\mathbf{e}_1\}. \tag{4}$$

Let's go with that assumption for a moment. Then our argument says to choose an arbitrary vector $\mathbf{u}_0 \in \ker(\ell)^\perp$. We could take $\mathbf{u}_0 = \mathbf{e}_2$ for example. Then a scaling

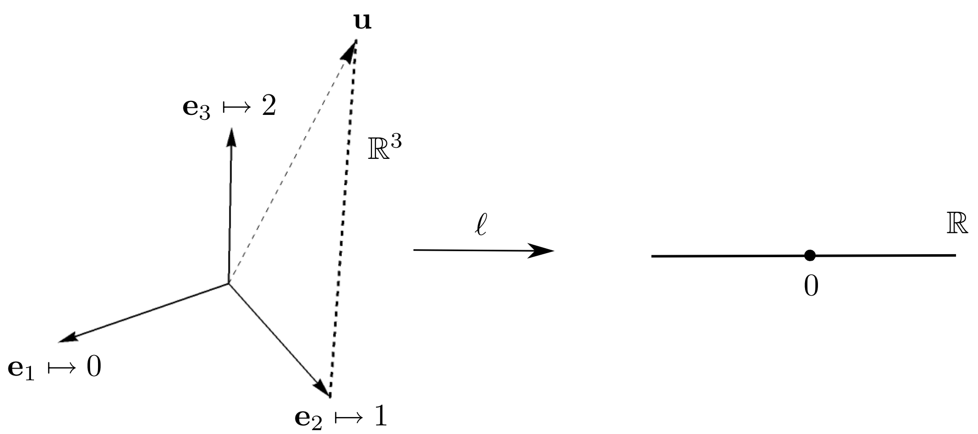

Figure 1: An example of a linear functional $\ell : \mathbb{R}^3 \to \mathbb{R}$.

$\mathbf{u} = \alpha \mathbf{u}_0$ is supposed to work. At this point, clearly something has gone wrong because we will never get $\mathbf{u} = (0, 1, 2)$ as a scaling of $\mathbf{u}_0 = (0, 1, 0)$. What has gone wrong?

What has gone wrong is that our identification of the null space in (4) is incorrect. The null space $\ker(\ell)$ is **larger** than $\text{span}\{\mathbf{e}_1\}$. In fact, if we had thought about it a little bit (and maybe you did) the dimension theorem says

$$\dim \text{Dom}(\ell) = \dim \text{Im}(\ell) + \dim \ker(\ell). \tag{5}$$

In this case, $\dim \ker(\ell) = \dim \text{Dom}(\ell) - \dim \text{Im}(\ell) = 3 - 1 = 2$. And in general, for $\ell : \mathbb{R}^n \to \mathbb{R}^1$, we must have $\dim \ker(\ell) = n - 1$. Thus, in the finite dimensional case, we must have a rather **large** kernel. In particular, the orthogonal complement of $\ker(\ell)$ must always satisfy

$$\dim \ker(\ell)^\perp = 1.$$

Thus, it is no surprise to find the representing vector there. This is the underlying idea also in the case of an infinite dimensional Hilbert space:

> Riesz representation follows because $\ker(L)$ is **large** to the extent that $\ker(L)^\perp$ is **one dimensional**.

We will return to this point. For the moment, let us give a more careful account of our example.

We should be able to find another vector in $\ker(\ell)$ for the linear map $\ell : \mathbb{R}^3 \to \mathbb{R}$ determined by (3). A moment's thought tells us $\ell : (0, -2, 1) \mapsto 0 \cdot 0 - 2 \cdot 1 + 1 \cdot 2 = 0$.

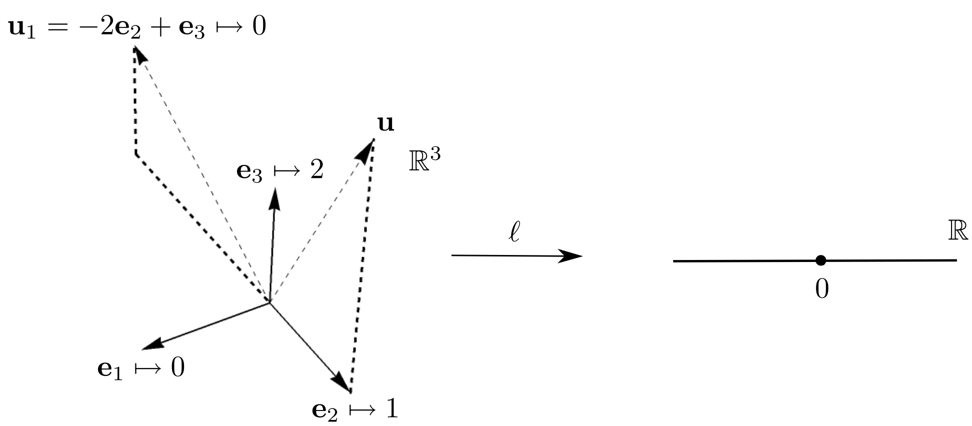

Figure 2: An example of a linear functional $\ell : \mathbb{R}^3 \to \mathbb{R}$; correctly identified null space.

Thus, $\ker(\ell)$ is a (two dimensional) plane and $\mathbf{u} = (0, 1, 2)$ is clearly a normal to that plane.

In the infinite dimensional case we do not have recourse to the dimension theorem/relation (5) telling us $\ker(\ell)$ is large, say of dimension $n - 1$, so that $\ker(\ell)^\perp$ is small, having dimension 1. These statements, however, do translate into infinite dimensions (for a continuous functional $L \in \mathcal{H}^*$ on a Hilbert space $\mathcal{H}$) in the following form

$\ker(L)$ is **large** to the extent that

$$\dim \ker(L)^\perp = 1 \qquad \text{and} \qquad \mathcal{H} = \ker(L) \oplus \ker(L)^\perp.$$

The crucial smallness of $\ker(L)^\perp$ holds in particular, so that representation is to be expected. I mention these things now, in part, because these facts do not come out in the proof. In the proof, we do something that uses much weaker hypotheses but is, in some sense, much trickier. Nevertheless, it is worth noting that this (true) description is what underlies and drives the result.

As just mentioned, we more or less need to do something rather tricky to show

$$\mathbf{u} \cdot \mathbf{v} = \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2} \mathbf{u}_0 \cdot \mathbf{v} = \ell(\mathbf{v}) \qquad \text{for all } \mathbf{v} \in \mathbb{R}^n$$

as required by the choice (2). Frigyes Riesz seems to have been particularly good at this kind of trickery. Given our previous discussion based on misidentification of

$\ker(\ell)$, it is clear that we must use rather strongly that $\mathbf{u}_0 \in \ker(\ell)^\perp$. The intuition is that we must use this fact in some way that takes account of the fact that $\mathbf{u}_0$ is orthogonal to **every** vector in $\ker(\ell)$, and not just some of them. In the finite dimensional case, every vector $\mathbf{v}$ decomposes uniqely as a sum

$$\mathbf{v} = \mathbf{z} + \mathbf{w} \qquad \text{for some } \mathbf{z} \in \ker(\ell) \text{ and some } \mathbf{w} \in \ker(\ell)^\perp.$$

In fact, taking $\ker(\ell)^\perp = \text{span}\{\mathbf{u}_0\}$, we can write

$$\mathbf{v} = \left(\mathbf{v} - \frac{\mathbf{v}\cdot\mathbf{u}_0}{|\mathbf{u}_0|^2}\mathbf{u}_0\right) + \frac{\mathbf{v}\cdot\mathbf{u}_0}{|\mathbf{u}_0|^2}\mathbf{u}_0 \tag{6}$$

where $\mathbf{w} = (\mathbf{v}\cdot\mathbf{u}_0)\mathbf{u}_0/|\mathbf{u}_0|^2$ is the projection of $\mathbf{v}$ onto $\ker(\ell)^\perp$. You may recognize this construction as a part of the Gram-Schmidt orthonormalization procedure, and it is clear that the residual vector

$$\mathbf{z} = \mathbf{v} - (\mathbf{v}\cdot\mathbf{u}_0)\mathbf{u}_0/|\mathbf{u}_0|^2 \tag{7}$$

is in $[\ker(\ell)^\perp]^\perp = \ker(\ell)$. What is immediately clear from the Gram-Schmidt construction is that $\mathbf{z} \in \text{span}\{\mathbf{u}_0\}^\perp = [\ker(\ell)^\perp]^\perp$. The fact that $[\ker(\ell)^\perp]^\perp = \ker(\ell)$ or more generally that the **double orthogonal complement** $V^{\perp\perp} = (V^\perp)^\perp$ satisfies

$$V^{\perp\perp} = V \qquad \text{for any subspace } V$$

requires proof.

Lemma 1 *If V is a subspace of $\mathbb{R}^n$, then $V^{\perp\perp} = V$.*

Proof: One inclusion is easy. Specifically, if $\mathbf{v} \in V$, then clearly $\mathbf{v}\cdot\mathbf{w} = 0$ for all $\mathbf{w} \in V^\perp$. This follows from the definition of

$$V^\perp = \{\mathbf{w} \in \mathbb{R}^n : \mathbf{w}\cdot\mathbf{v} = 0 \text{ for all } \mathbf{v} \in V\}.$$

But this is also the definition of what it means to have $\mathbf{v} \in (V^\perp)^\perp = V^{\perp\perp}$.

The reverse inclusion is trickier: If $\mathbf{v} \in V^{\perp\perp}$, then there are unique vetors $\mathbf{z} \in V$ and $\mathbf{w} \in V^\perp$ with

$$\mathbf{v} = \mathbf{z} + \mathbf{w}. \tag{8}$$

Recalling from the first inclusion that $V \subset V^{\perp\perp}$, we know $\mathbf{z} \in V^{\perp\perp}$ and, of course, $V^{\perp\perp}$ is a subspace. Therefore,

$$\mathbf{w} = \mathbf{v} - \mathbf{z} \in V^{\perp\perp} \cap V^\perp.$$

This implies $|\mathbf{w}|^2 = \mathbf{w}\cdot\mathbf{w} = 0$. Hence $\mathbf{w} = 0$ and $\mathbf{v} = \mathbf{z} \in V$. $\square$

It will be noted that we have used the following result to obtain (8).

Lemma 2 *If V is a subspace of $\mathbb{R}^n$, then $\mathbb{R}^n = V \oplus V^\perp$. That is, each $\mathbf{v} \in \mathbb{R}^n$ is expressed uniquely as $\mathbf{v} = \mathbf{z} + \mathbf{w}$ with $\mathbf{z} \in V$ and $\mathbf{w} \in V^\perp$. The vector $\mathbf{z}$ is called the* **projection** *of $\mathbf{v}$ onto V.*

Proof: If we want to get out of this easily, then we must allow recourse to the fact that $\mathbb{R}^n$ and V are finite dimensional, i.e., these spaces admit bases with finitely many elements. Say $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis for V which we can assume (by Gram-Schmidt orthonormalization) is an orthonormal basis. Then

$$\mathbf{z} = \mathrm{proj}_V(\mathbf{v}) = \textstyle\sum_{j=1}^k (\mathbf{v} \cdot \mathbf{u}_j)\, \mathbf{u}_j.$$

As a direct extension of our assertion concerning (6) we see that $\mathbf{v} - \mathbf{z} \in V^\perp$. Thus, we have $\mathbf{v} = \mathbf{z} + \mathbf{w}$ with $\mathbf{z} \in V$ and $\mathbf{w} \in V^\perp$.

For uniqueness, note that if $\mathbf{v} = \tilde{\mathbf{z}} + \tilde{\mathbf{w}} = \mathbf{z} + \mathbf{w}$ with $\tilde{\mathbf{z}} \in V$ and $\tilde{\mathbf{w}} \in V^\perp$, then $\mathbf{z} - \tilde{\mathbf{z}} = \tilde{\mathbf{w}} - \mathbf{w} \in V \cap V^\perp$. That is,

$$|\tilde{\mathbf{w}} - \mathbf{w}|^2 = (\tilde{\mathbf{w}} - \mathbf{w}) \cdot (\tilde{\mathbf{w}} - \mathbf{w}) = 0 \qquad \text{and} \qquad |\mathbf{z} - \tilde{\mathbf{z}}|^2 = 0. \qquad \square$$

Returning to our discussion of (6) and (7), if we know the vector $\mathbf{z}$ from (7) satisfies $\mathbf{z} \in \ker(\ell)$, then it is immediate that

$$\ell(\mathbf{v}) = \ell\left(\frac{\mathbf{v} \cdot \mathbf{u}_0}{|\mathbf{u}_0|^2}\, \mathbf{u}_0\right) = \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\, \mathbf{v} \cdot \mathbf{u}_0 = \mathbf{v} \cdot \frac{\ell(\mathbf{u}_0)}{|\mathbf{u}_0|^2}\, \mathbf{u}_0,$$

and we have the Riesz representation proposed in connection with (2).

The argument just given using the finite orthonormal basis for V does not work when V is infinite dimensional. Nevertheless, the assertions of Lemmas 1 and 2 do both hold for any **closed subspace** V **of a Hilbert space**. In fact, once Lemma 2 is established for a closed subspace V of a Hilbert space, then the argument given for Lemma 1 is valid in the same context. The following exercise, establishing the existence of projections in a Hilbert space, is the key.

Exercise 1 *Let V be a closed subspace of a Hilbert space $\mathcal{H}$ and let $p \in \mathcal{H} \backslash V$. Consider a sequence of points $x_j \in V$ for $j = 1, 2, 3, \dots$ with*

$$\lim_{j \to \infty} \|x_j = h\|_{\mathcal{H}} = \mathrm{dist}(p, V) = \inf_{x \in V} \|x - p\|_{\mathcal{H}}. \tag{9}$$

(i) *Show the sequence $\{x_j\}_{j=1}^\infty$ is Cauchy.*

(ii) *Use the completeness of $\mathcal{H}$ (and the fact that V is closed) to conclude the limit $\lim_{j \to \infty} x_j = x$ achieves the minimum value in the infemum of (9).*

(iii) *Show the difference $x - p \in V^\perp$.*