Riesz Representation Theorem – a simple version

P. Reany

July 25, 2024

Abstract

When I first saw a proof of this theorem, I was confused by it, first by the formalism, and second by the fact that the solution was appeerently 'pull out of the blue'. I will attempt to address both of these issues. The point of this theorem is to show that we can represent a linear map of vectors to scalars by means of an inner product.

1 The Theorem

Let V be a finite-dimensional vector space over a field F. Let ϕ be a linear map from V to F,

$$\phi(v) \in F\,. \tag{1}$$

For simplicity's sake, I make the assumption that the field is the real numbers. Then, for all $u \in V$, there exists a specific $v \in V$ such that

$$\phi(u) = \langle u, v \rangle\,, \tag{2}$$

where $\langle u, v \rangle$ is the standard inner product on vectors.[1] The functional[2] ϕ is said to be an element of the vector space *dual* to V and it is represented as V^*.

Okay, where did this vector v come from? Seemingly, from out of the blue. Maybe the insight needed to solve for v is to look at some other objects or aspects of this vector space that can be inferred from 'out of the blue'.

Here's the heuristic I'm suggesting: Let S be a system on which we need to prove some theorem (or conjecture) T . If what we need to show requires us to pull something 'out of the blue', maybe we should look at known objects and/or aspects of S that can be inferred 'from out the blue'.

2 The Proof

So, following this line of reasoning, what can we infer about finifte vector space V? First, we can infer that it has a basis $B = \{e_1, e_2, \ldots, e_n\}$. We know this because of the proved theorem that *every vector space has a basis*. Furthermore, since V is finite-dimensional, we can employ the *Gram–Schmidt process* to convert B into an orthomormal basis, if it wasn't already. So, let's assume that B is orthonormal. In more mathematical terms, we have that

$$\langle e_i, e_j \rangle = \delta_{ij}\,, \tag{3}$$

[1] There is a far more involved form of this theorem that is stated and proved from the domain of functional analysis.
[2] A functional is a function that maps vectors to scalars (the field elements).

where δ_{ij} is the Kronecker delta.

So, what can we say now?

$$\phi(e_i) \in F \quad \text{for all } i \in [1..n] \,. \tag{4}$$

We don't really care what these scalars are – just that they exist. Next, let's expand an arbitrary vector $u \in V$ in the basis B:

$$u = \sum_{i=1}^{n} u_i e_i = u_i e_i \quad (\text{where } u_i \equiv \langle u, e_i \rangle) \,, \tag{5}$$

where I am using the *Einstein summation convention* that repeated indices in a given term mean we will automatically perform the summation on them (unless stated otherwise).[3] Then

$$\phi(u) = \phi(u_i e_i) = u_i \phi(e_i) \,, \tag{6}$$

where we have used the linearity of ϕ.

Here's where we employ the hypothesis of a genetic connection between these two 'out of the blue' objects. Since we have n scalars $\phi(e_i)$ sitting around not doing very much, let's just define a special vector out of them – 'out of the blue', beginning with the vector's components:

$$v_i \equiv \phi(e_i) \quad \text{for all } i \in [1..n] \,. \tag{7}$$

Now that we have the components, we can construct the vector v:

$$v \equiv v_i e_i \,. \tag{8}$$

Okay, let's take the inner product of u and v and see where we end up:[4]

$$\begin{aligned} \langle u, v \rangle &= \langle u_i e_i, v_j e_j \rangle = u_i v_j \langle e_i, e_j \rangle \\ &= u_i v_j \delta_{ij} = u_i v_i \\ &= u_i \phi(e_i) = \phi(u_i e_i) \\ &= \phi(u) \,. \end{aligned} \tag{9}$$

So, we've proved what we needed to prove. But what about the uniqueness of v? Could some other vector, v' do just as well? If it could then we need to allow the following:

$$\langle u, v \rangle = \langle u, v' \rangle \,. \tag{10}$$

But this is true if and only if

$$\langle u, v \rangle - \langle u, v' \rangle = 0 \quad \text{if and only if} \quad \langle u, v - v' \rangle = 0 \,. \tag{11}$$

So, we have two choices. Either $v - v' = 0$ or $v - v'$ is orthogonal to u. But we cannot have this vector difference to depend on u, because it's already strictly determined by ϕ and applies to all u. Therefore we conclude that $v = v'$, and thus v is unique.

A second rationale is the following: Since $\langle u, v - v' \rangle = 0$ must be true for all u, it must be true for $u = v - v'$, in which case we have then that

$$\langle v - v', v - v' \rangle = |\, v - v' \,|^2 = 0 \,. \tag{12}$$

But this is true if and only if $v - v' = 0$ or $v = v'$, so v is unique.

[3] In more rigorous treatments of the subject, we would have to be concerned about using both upper and lower indices.

[4] We must expand these vectors using different indicies to get the right answer, so that the two summations can be made independently of each other.