LINEAR FUNCTIONALS AND ADJOINT OPERATORS

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog.

References: edX online course MIT 8.05.1x Week 4.

Sheldon Axler (2015), *Linear Algebra Done Right*, 3rd edition, Springer. Chapters 3.F, 7.

Linear functionals. A *linear functional* is a linear map $\phi(v)$ from a vector space V to the number field $\mathbb{F}$ which satisfies the two properties

(1) $\phi(v_1+v_2)=\phi(v_1)+\phi(v_2)$, with $v_1,v_2\in V$.
(2) $\phi(av)=a\phi(v)$ for all $v\in V$ and $a\in\mathbb{F}$.

That is, a linear functional acts on a vector and produces a number as output.

A linear functional is actually a vector space, since it satisfies all the required axioms. Many of these axioms are satisfied because V on which ϕ acts is a vector space. The only axiom that requires a bit of examination is the existence of an additive identity. This requires $\phi(v+0)=\phi(v)$. From property 1 above, this means that $\phi(0)=0$.

From the definition above, we can prove that any linear functional can be written as an inner product.

Theorem 1. *For any linear functional ϕ on V there is a unique vector $u\in V$ such that $\phi(v)=\langle u,v\rangle$ for all $v\in V$.*

Proof. We can write any vector v in terms of an orthonormal basis $(e_1,\ldots,e_n)$ as

$$v=\sum_{i=1}^{n}\langle e_i,v\rangle e_i \tag{1}$$

By applying the two properties of a linear functional above, we have

$$\phi(v) = \phi\left(\sum_{i=1}^{n}\langle e_i, v\rangle e_i\right) \tag{2}$$

$$= \sum_{i=1}^{n}\langle e_i, v\rangle \phi(e_i) \tag{3}$$

$$= \sum_{i=1}^{n}\langle e_i, \phi(e_i) v\rangle \tag{4}$$

$$= \sum_{i=1}^{n}\left\langle \phi(e_i)^* e_i, v\right\rangle \tag{5}$$

$$= \left\langle \left[\sum_{i=1}^{n}\phi(e_i)^* e_i\right], v\right\rangle \tag{6}$$

$$\equiv \langle u, v\rangle \tag{7}$$

where

$$u \equiv \sum_{i=1}^{n}\phi(e_i)^* e_i \tag{8}$$

We were able to move $\phi(e_i)$ inside the inner product in the third line above since$\phi(e_i)$ is just a number.

To prove that u is unique, as usual we suppose there is another u' that gives the same result as u for all $v \in V$. This means that $\langle u', v\rangle = \langle u, v\rangle$ or $\langle u' - u, v\rangle = 0$ for all v. We can then choose $v = u' - u$, giving $\langle u' - u, u' - u\rangle = 0$, which implies that $u' - u = 0$ so $u' = u$. $\square$

Adjoint operators. Suppose we have some linear operator T and some fixed vector u. We can then form the inner product

$$\phi(v) = \langle u, Tv\rangle \tag{9}$$

$\phi(v)$ is a linear functional since it satisfies the two properties specified earlier. It is now stated in Zwiebach's notes that because ϕ is a linear functional, we can write it in the form $\langle w, v\rangle$ for some vector w. It's not clear to me that this follows directly, since the original definition of a linear functional applied to the entire vector space V, whereas here we don't know whether T is a surjective operator, that is, whether the range of Tv is the entire space V. The motivation behind this step is the definition of the adjoint operator, but in Axler's book (chapter 7.A), an adjoint is just defined directly without any motivation from linear functionals.

Anyway, we'll just go with Zwiebach's argument, since the rest of the derivation is fairly easy to follow. We assume that for a suitable vector w we have

$$\langle u, Tv \rangle = \langle w, v \rangle \tag{10}$$

The vector w depends on both the operator T and the vector u, so we can write it as a function of u, using the notation

$$w = T^\dagger u \tag{11}$$

This gives us the relation

$$\langle u, Tv \rangle = \left\langle T^\dagger u, v \right\rangle \tag{12}$$

At this stage, we can't be sure that $T^\dagger$ is a linear operator; it may be some non-linear map from one vector to another. However, we have

Theorem 2. *The operator $T^\dagger$, called the adjoint of T, is a linear operator: $T^\dagger \in \mathcal{L}(V)$.*

Proof. Consider

$$\langle u_1 + u_2, Tv \rangle = \left\langle T^\dagger (u_1 + u_2), v \right\rangle \tag{13}$$

$$\langle u_1 + u_2, Tv \rangle = \langle u_1, Tv \rangle + \langle u_2, Tv \rangle \tag{14}$$

$$= \left\langle T^\dagger u_1, v \right\rangle + \left\langle T^\dagger u_2, v \right\rangle \tag{15}$$

$$= \left\langle T^\dagger u_1 + T^\dagger u_2, v \right\rangle \tag{16}$$

Comparing the first and last lines gives us

$$T^\dagger (u_1 + u_2) = T^\dagger u_1 + T^\dagger u_2 \tag{17}$$

A similar argument can be used for multiplication by a number:

$$\langle au, Tv \rangle = \left\langle T^\dagger (au), v \right\rangle \tag{18}$$

$$\langle au, Tv \rangle = a^* \langle u, Tv \rangle \tag{19}$$

$$= a^* \left\langle T^\dagger u, v \right\rangle \tag{20}$$

$$= \left\langle a T^\dagger u, v \right\rangle \tag{21}$$

Again, comparing the first and last lines we have

$$T^\dagger (au) = a T^\dagger u \tag{22}$$

Thus $T^{\dagger}$ satisfies the two conditions required for linearity. $\square$

A couple of other results follow fairly easily (proofs are in Zwiebach's notes, if you're interested):

$$\begin{aligned}(ST)^{\dagger} &= T^{\dagger}S^{\dagger} &(23)\\ \left(S^{\dagger}\right)^{\dagger} &= S &(24)\end{aligned}$$

A very important result is the representation of adjoint operators in matrix form.

If we have an orthonormal basis $(e_1,\ldots,e_n)$ and an operator T, then T transforms the basis according to (using the summation convention):

$$Te_k = T_{ik}e_i \tag{25}$$

Thus the matrix elements in this basis are found by taking the inner product with e_j:

$$\langle e_j, T_{ik}e_i\rangle = T_{ik}\langle e_j, e_i\rangle = T_{ik}\delta_{ji} = T_{jk} \tag{26}$$

Also,

$$T^{\dagger}e_j = T^{\dagger}_{ij}e_i \tag{27}$$

Taking the inner product on the right with e_k we get

$$\left\langle T^{\dagger}e_j, e_k\right\rangle = \left\langle T^{\dagger}_{ij}e_i, e_k\right\rangle = \left(T^{\dagger}_{ij}\right)^{*}\langle e_i, e_k\rangle = \left(T^{\dagger}\right)^{*}_{kj} \tag{28}$$

We can take the T_{ik} and $T^{\dagger}_{ik}$ outside the inner product as they are just numbers. Using this, we have

$$\begin{aligned}\left\langle T^{\dagger}e_j, e_k\right\rangle &= \langle e_j, T_{ik}e_i\rangle &(29)\\ \left(T^{\dagger}\right)^{*}_{kj} &= T_{jk} &(30)\end{aligned}$$

That is, in an orthonormal basis, the adjoint matrix is the complex conjugate transpose of the original matrix:

$$T^{\dagger} = (T^{*})^{T} \tag{31}$$

The superscript T indicates 'transpose', not another operator!

PINGBACKS

Pingback: Unitary operators