The matter of dual space and dual basis. PhL 8.15.15

I think this is something I already know about from tensor doc. Recall there from page 75:

[pages 82 and 83]

Notes on Reciprocity (Duality)

1. Suppose some set of vectors $\mathbf{b}_n$ forms a complete basis for x-space. Can one find a set of vectors $\mathbf{B}_n$ that have the property

$$\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n} \; ? \qquad \text{// duality relation; } \mathbf{B}_a \text{ and } \mathbf{b}_a \text{ are reciprocal} \qquad (6.2.8)$$

As shown below, the answer is normally "yes", and the vectors $\mathbf{B}_n$ are uniquely determined by the $\mathbf{b}_n$. One says that the set $\{\mathbf{B}_n\}$ is "dual to" the set $\{\mathbf{b}_n\}$ and vice versa. If we regard $\mathbf{b}_n$ as a basis, then $\mathbf{B}_n$ is the "dual basis", and $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$ is the "duality relation". Another terminology is that the vectors $\mathbf{B}_n$ are "reciprocal to" the vectors $\mathbf{b}_n$ and vice versa.

3. One can solve for the $\mathbf{B}_n$ in terms of the $\mathbf{b}_n$. Each $\mathbf{B}_n$ has N components, so there are N^2 unknowns. The duality relation $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$ is a set of N^2 equations. This is basically a Cramer's Rule problem in N^2 variables. Since the $\mathbf{b}_n$ form a complete basis, one can expand $\mathbf{B}_m$ on the $\mathbf{b}_n$ with some coefficients we will call w'_{mn} (at this point w'_{mn} is unknown),

$$\mathbf{B}_m = w'_{mn}\mathbf{b}_n \; . \qquad (6.2.9)$$

Then from (6.2.8),

$$\delta_{m,k} = \mathbf{B}_m \bullet \mathbf{b}_k = w'_{mn}\, \mathbf{b}_n \bullet \mathbf{b}_k \; . \qquad // \quad \mathbf{b}_n \bullet \mathbf{b}_k = \bar{g}_{ij}(b_n)_i(b_k)_j \; . \qquad (6.2.10)$$

Define matrix W' by,

$$W'_{nk} \equiv \mathbf{b}_n \bullet \mathbf{b}_k \qquad (6.2.11)$$

and note that W'_{nk} is symmetric. Then (6.2.10) says

$$\delta_{m,k} = w'_{mn}W'_{nk} \quad \text{or} \quad w'W' = 1 \quad \text{or} \quad w' = W'^{-1} \; . \qquad (6.2.12)$$

Assuming for the moment that $\det W' \neq 0$, the solution is given by $w' = W'^{-1}$. The (5.6.4) "Digression" showed that $(A^{-1})^T = (A^T)^{-1}$ for invertible A, so $w'^T = (W'^{-1})^T = (W'^T)^{-1} = W'^{-1} = w'$ and therefore w' is symmetric as well. Since W' is known from (6.2.11), $w' = W'^{-1}$ and the $\mathbf{B}_m = w'_{mn}\mathbf{b}_n$ of (6.2.9) have been found. Finally,

$$\mathbf{B}_m \bullet \mathbf{B}_n = w'_{mi}\mathbf{b}_i \bullet w'_{nj}\mathbf{b}_j = w'_{mi}w'_{nj}\, \mathbf{b}_i \bullet \mathbf{b}_j = w'_{mi}w'_{nj}W'_{ji} = w'_{mi}\delta_{ni} = w'_{nm} \; . \qquad (6.2.13)$$

So in this tensor doc context, I can regard the $\mathbf{b}_i$ = Sja $\mathbf{v}_i$ as some basis vectors in the vector space spanned by the $\mathbf{b}_i$, and then I can regard the $\mathbf{B}_i$ = Sja $\boldsymbol{\lambda}_i$ as a set of row vectors that span the "row space".

You can think of each row vector application as a *linear functional* (which is a special case of a linear map)

$\mathbf{B}_m \bullet \mathbf{x}$ = number $\qquad B_m : \ R^n \rightarrow R \qquad \mathbf{x}$ in R^n

$\boldsymbol{\lambda}_m \bullet \mathbf{x}$ = number $\qquad \lambda_m : \ R^n \rightarrow R$

In fact any row vector acts as a functional in this sense, not just the basis vectors. Fine.

$\mathbf{y} \bullet \mathbf{x}$ = number $\qquad y : \ R^n \rightarrow R \qquad \mathbf{y}$ in $R^{n*} \qquad \mathbf{x}$ in R^n

Sja wants to refer to the vector space of column vectors $\mathbf{x}$ as V = subset of R^n spanned by the $\mathbf{b}_j = \mathbf{e}_j$.

Sja wants to refer to the vector space of row vectors $\mathbf{y}$ as V* = subset of R^{n*} spanned by the $\mathbf{b}_j = \mathbf{e}_j$.

Then the space V* is "dual to" V, and V* is "the dual space". This all seems quite innocent.

[page 84]

Example 7.2 . Here V is a space of real functions (not just a piece of R^n) in the Stakgold sense defined on [a,b]. The functional of interest is $\mu[f] = \int f\,dx$ which gives you a real number. Somehow I*f = integral value. Sja is only here giving an example of a linear functional, he says nothing any "dual space" to V.
[below I show that this particular functional is just one of a infinite number of such in V*]

Example 7.3. In this example, the functional is $\mu[x] = \mathbf{v} \bullet \mathbf{x}$ for some fixed $\mathbf{v}$ where now we have the "standard inner product" used in R^n. I have been using this notation all along, while Sja has been avoiding it.

Now, suppose $\mathbf{v}_i$ are the basis vectors of V of dimension n. ($\mathbf{b}_i$ in tensor doc). Sja expands a general vector $\mathbf{v}$ as

$$\mathbf{v} = \Sigma_j \, c_j \, \mathbf{v}_j \quad = \quad \Sigma_j \, \lambda_j(\mathbf{v}) \, \mathbf{v}_j \qquad\qquad c_j \ = \ \lambda_j(\mathbf{v})$$

So this object $\lambda_j(\mathbf{v})$ is a row vector of constants c_j which when linearly combined with the v_j gives some vector $\mathbf{v}$. I would say then that

$$\lambda_j(\mathbf{v}) \ = \mathbf{v} \bullet \mathbf{v}_j \qquad \text{and then} \qquad \mathbf{v} = \Sigma_j \, (\mathbf{v} \bullet \mathbf{v}_j) \, \mathbf{v}_j$$

But I think his definition does not require the existence of an inner product, it just requires a vector space. Working again with his definition, it seems clear that

$$\lambda_j(\mathbf{v}_j) = 1$$
$$\lambda_j(\mathbf{v}_i) = 0 \qquad i \neq j$$

and you arrive at this just looking at the linear combination. So we then have this set of coordinate functions λ_i and each of these is a functional mapping $R^n \rightarrow R$ such as $\lambda_j(\mathbf{v}) = c_j$.

OK, so I think the idea is to try to avoid talking about dot products to be general to any vector space, not just a Hilbert space.

Wiki and Stakgold

I looked at wiki at this point. It says that the space V* is the space of all linear functionals on V. So let's look again at the above examples. When $V = R^n$ which is a finite dimensional space, I think this is the most general linear functional you can write:

$$\mathbf{v} = \Sigma_{j=1}^{n} c_j \mathbf{v}_j = \Sigma_j \lambda_j(\mathbf{v}) \mathbf{v}_j \qquad c_j = \lambda_j(\mathbf{v})$$

Now exactly what is the linear functional here? I would say that *each* of the λ_j functions is a linear functional. I show linearity below. So for each j we have λ_j: $R^n \rightarrow R$. Thus, we are aware right away of a set of n different functionals on V. Below we show that these λ_j functionals form a basis for V* and therefore V* has dimension n. Wiki would say that the elements of V* are called covectors, so therefore each of the λ_j functions would be a covector. Wiki allows that $\lambda_j(\mathbf{v}) = [\lambda_j,\mathbf{v}] = \langle\lambda_j,\mathbf{v}\rangle$ are other common notations.

Wiki says for finite n V, you can take e_i as the V basis and then e^i is the V* basis where $e^i e_j = \delta_{ij}$, so this is the same as my tensor doc idea. The dual space is the space of row vectors. Wiki goes on to talk about V being ∞ dimensional, but does not really address any space of functions. But:

Suppose V instead of being R^n is some infinite dimensional function space. Consider $\mu[f] = \int f\,dx$ where f lies in V. Here μ is a linear functional and we have μ: $V \rightarrow R$. So what is the dual space V* in this case? I have displayed only one linear functional, but V* is the space of all such. So I don't have an answer here. Notice that $\mu[f] = \int f^{\alpha}\,dx$ for $\alpha \neq 1$ would not be a *linear* functional. $\mu[f] = \int f\,dx\,s(x)$ might also be a linear functional where s(x) is a weight function. Stak talks about this subject.

In Stakgold, a functional is called T[**x**] where **x** lies in V. When V is a Hilbert Space, you can state the famous Riesz Representation Theorem Stak chap 2.

" Any continuous and linear functional which acts on any S in any Hilbert space V can be "represented" as T[x] = <x,f> (for all x in V) where f is some unique element of V "

Here we have **x** being some vector in V, and the theorem claims that any linear functional can be expressed as the inner product of **x** with some other unique element of V which above is called f. Thus, the number of linear functionals would be the same as the number vectors **x** in V. For $V = R^n$ both these numbers are infinite, and you can think of f as being the row vectors. For a function space where x is a function, I might restate this theorem as

" Any continuous and linear functional which acts on any S in any function Hilbert space V can be "represented" as T[f] = <f,F> (for all f in V) where F is some unique function of V " . Then maybe

$$T[f] = \int f(x)\, F(x)\, dx$$

Then the example above happens to use the function F(x) = 1. So then my conjecture of weight functions would be correct.

Note that Riesz is in the context of a Hilbert Space, not a simple vector space.

Now back to Sjamaar:

Lemma 7.4. The λ_i form a basis for V*.

Now what does λ_i mean without an argument? It is a *function* so can only exist in a space of functions. He would have then to show that if λ were an arbitrary function, you could write $\lambda = \Sigma\, d_i\, \lambda_i$ where d_i are some constant coefficients. If we could construct d_i, then maybe we have our proof. So evaluate this arbitrary function λ for one of the V-space basis vectors

$$\lambda(\mathbf{v}_j) = \Sigma\, d_i\, \lambda_i(\mathbf{v}_j) = \Sigma\, d_i\, \delta_{i,j} = d_j$$

So there you have d_j constructed. We then have

$$\lambda = \Sigma\, \lambda(\mathbf{v}_j)\, \lambda_i$$

and we have then shown that λ_i is a basis since it can generate any function λ by linear combination. So I guess this looks good, and again we have avoided mention of any inner product. If you had an inner product, you could write $\lambda_j(\mathbf{v}) = \mathbf{v} \bullet \mathbf{v}_j$ but we don't want to assume an inner product, so we just have these functions λ_j. Fine by me.

Because the set of n functions λ_i spans V*, we must have dim(V*) = n.

Fact : The V* basis of *functions* $\{\lambda_i\}$ is dual to the basis of *vectors* $\{\mathbf{v}_j\}$ of V.

Note that λ_i is a scalar function of a vector just as $\lambda_j(\mathbf{v}) = \mathbf{v} \bullet \mathbf{v}_j$ would suggest were there an IP. Note that the functions λ_i are in fact functionals since they map to R.

Are the λ_i *linear* functionals? Let's investigate:

$\mathbf{v} = \Sigma_j \lambda_j(\mathbf{v})\, \mathbf{v}_j$

$(\alpha\mathbf{v}) = \Sigma_j \lambda_j(\alpha\mathbf{v})\, \mathbf{v}_j = \alpha\, [\, \Sigma_j \lambda_j(\mathbf{v})\, \mathbf{v}_j\,] = \Sigma_j\, (\alpha\, \lambda_j(\mathbf{v}))\, \mathbf{v}_j = \Sigma_j\, (\lambda_j(\alpha\mathbf{v}))\, \mathbf{v}_j$

Since the coefficients are unique, we have

$\lambda_j(\alpha\mathbf{v}) = \alpha\, \lambda_j(\mathbf{v})$ first part of showing linear.

Next,

$\mathbf{v} + \mathbf{v}' = \Sigma_j\, [\lambda_j(\mathbf{v}) + \lambda_j(\mathbf{v}')]\, \mathbf{v}_j = \Sigma_j \lambda_j(\mathbf{v}+\mathbf{v}')\, \mathbf{v}_j \quad \Rightarrow$

$\lambda_j(\mathbf{v}+\mathbf{v}') = \lambda_j(\mathbf{v}) + \lambda_j(\mathbf{v}')$

So YES, the functions λ_i are linear functions and are linear functionals. No dot product needed.

Example 7.5. Let $\mathbf{e}_i$ be the basis vectors of V, as usual. We know that $dx_i = \mathbf{e}_i^T$ which lies in the dual space R^{n*} . It seems pretty reasonable that the $\{\mathbf{e}_i^T\}$ form a basis of R^{n*} and in fact

$\mathbf{e}_i^T \mathbf{e}_j = e_i \bullet e_j = \delta_{i,j}$

so then we can say: " the set $\{dx_i\}$ form a basis in R^{n*}". I guess one could write

$\mathbf{v} = \Sigma\, v_i \mathbf{e}_i$ general vector in R^n
$v^T = \Sigma\, v_i \mathbf{e}_i^T$ general vector in R^{n*}

So in this example you can really associate the dual space with simple row vectors, not so much abstract functions.

Suppose we have L: V → W where V has basis $\mathbf{v}_i$ and W has basis $\mathbf{w}_i$. Consider

Lv_j = some vector in W $= \Sigma_i\, \ell_{i,j}\, \mathbf{w}_i$ defines some numbers $\ell_{i,j}$

Again, no inner product yet.

Lemma 7.6. Imagine this situation for L: V → W

$\mathbf{v}_i$ = basis in V
λ_i = basis for V*

$\mathbf{w}_i$ = basis for W
μ_i = basis for W*

Then compute

$$\mu_i(Lv_j) = \mu_i(\Sigma_k \ell_{k,j} w_k) = \Sigma_k \ell_{k,j} \mu_i(w_k) = \Sigma_k \ell_{k,j} \delta_{i,k} = \ell_{i,j}$$

Not sure why this is interesting, but it is correct.

[page 85]

Definition: Sja defines the simple property of a **k-multilinear function** $\lambda(\mathbf{v_1},\mathbf{v_2}.....\mathbf{v_k})$ This is $\lambda : V^k \rightarrow R$, so there are k distinct vector arguments. If there are two arguments, it is bilinear = 2-multilinear. Some simple examples are given. [these are in fact multilinear functionals since maps to R]

Clarification: The vectors $\mathbf{v_i}$ could lie in some space $V = R^n$ where n is unrelated to k. In general $\mathbf{v_i}$ in V. Do not assume that k = n, that is not implied!

Comment: I think this is the obvious generalization of a linear functional of one vector variable to many vector variables in the same space V. Of course $V^k = V \otimes V \otimes .. \otimes V$ which is a direct product space as in my tensor doc discussion.

Definition: If you add the property that swapping any two arguments negates λ, then you have an **alternating or asymmetric** k-multilinear function.

The big example here is $\det(\mathbf{c_1},\mathbf{c_2}....\mathbf{c_k})$ is a well-known alternating k-multilinear function. Sja has discussed this animal earlier, and there I broke it down into two properties, each of which determinants have. He gives a second simpler alternating example that is bilinear.

Now Sja does a very weird thing. He defines a single function which has the following name:

$$"\lambda_1\lambda_2....\lambda_k"$$

The arguments of this function are going to be $\mathbf{v_1},\mathbf{v_2}...\mathbf{v_k}$ which are the basis vectors of V. And here is how the function is defined:

$$\lambda_1\lambda_2....\lambda_k(\mathbf{v_1},\mathbf{v_2}...\mathbf{v_k}) \equiv \det[\lambda_i(\mathbf{v_j})]$$

The matrix here is this:

$$\begin{matrix} \lambda_1(\mathbf{v_1}) & \lambda_1(\mathbf{v_2}) & \lambda_1(\mathbf{v_3}) & ..... & \lambda_1(\mathbf{v_k}) \\ \lambda_2(\mathbf{v_1}) & \lambda_2(\mathbf{v_2}) & \lambda_2(\mathbf{v_3}) & ..... & \lambda_2(\mathbf{v_k}) \\ ... & & & & \\ \lambda_k(\mathbf{v_1}) & \lambda_k(\mathbf{v_2}) & \lambda_k(\mathbf{v_3}) & ..... & \lambda_k(\mathbf{v_k}) \end{matrix}$$

This fancy function is a k-multilinear function because that is the way determinants work. It is also alternating since THAT is also how they work.

This function $\lambda_1\lambda_2....\lambda_k$ is called "the **wedge product** of the λ's" function and one writes then

$$[\lambda_1 \wedge \lambda_2 \wedge....\wedge \lambda_k](\mathbf{v}_1,\mathbf{v}_2...\mathbf{v}_k) \equiv \det[\lambda_i(\mathbf{v}_j)]$$

I see the mechanical definitions, but I have *no idea* how this relates to anything I know about forms.

Question: Can we identify $\lambda_i(\mathbf{v}_j) = \delta_{i,j}$?
Answer: Well, in the current context, the $\mathbf{v}_i$ are arbitrary arguments of the multilinear function and are general vectors in V. They could be, but are in general *not*, the set of basis vectors. Only if we happened to evaluate the wedge thing at the basis vectors would be make such a claim.

A^kV = set of alternating k-multilinear functions.

This A^kV is in fact a vector space, and as such, it should have some basis.

Comment: Recall that V* is the collection of linear functionals on V -- the dual space. For finite V with basis *vectors* $\mathbf{v}_i$ we found that this dual space is spanned by the set of *functionals* λ_i. Each of these functionals is a function of **v** where **v** lies in V.

Now we generalize so that V is replaced by $V^k = V \otimes V \otimes .. \otimes V$. What is the corresponding dual space? Perhaps it is the set of *multilinear* functionals on V^k. I don't know how many there are right now, but write *one* of these as $\mu(\mathbf{v}_1, \mathbf{v}_2.....\mathbf{v}_k)$. I presume then that the set of all possible such multilinear functionals would be the dual space $(V^k)^*$. That would seem to be the logical generalization.

For reasons not yet clear, Sja wants us to consider only multilinear functionals μ which have the *alternating property*, and instead of $(V^k)^*$ he refers to this restricted space as AV^k. So before the reasons for being interested in sort of dual space are given, Sja wants to develop some properties of this space, and only then can he explain the motivation.

Sja is not referring to either $(V^k)^*$ or AV^k as a "dual space", but either seems a reasonable generalization of the dual space concept.

Naively you might think that

$$(V^k)^* = V^* \otimes V^* \otimes .. \otimes V^*$$

and then maybe the basis vectors of $(V^k)^*$ would be

$$\lambda_i \otimes \lambda_j \otimes.....\otimes \lambda_k \qquad\qquad i = 1..k,\ j = 1..k \text{ etc}$$

I will come back to this idea later. This approach however does not work in that alternating property. But maybe this would be the normal dual space of $(V^k)^*$. I think it is.

Properties of the space AV^k.

1. It is a vector space. Well, $(V^k)^*$ is a vector space since it is a dual space, and if restrict to multilinear functionals which have the alternating property, you still get a vector space. Key fact is that if you add two alternators, the sum will be alternating.

2. As a vector space, AV^k must have some basis vectors and must have some dimension.

[page 86]

Define I as our usual increasing multi index.

Example 7.11.

Here $V = R^3$ with the usual e_i basis vectors and the usual e_i^T dual space basis vectors.