Groups, Fields, and Vector Spaces

Homework #3 (2024-2025), Answers

Q1: Duals of infinite-dimensional spaces

A. Consider the (infinite-dimensional) vector space V of real-valued functions $f(x)$ on the interval $[0,1]$ *that are "nice" – continuous, smooth, integrable. For any $g \in V$, there is a mapping Ag from V to the base field, defined by $(Ag)(f) = \int_0^1 g(x)f(x)dx$. Show that $Ag \in V^*$, i.e., that it is a linear map from V to the base field.*

We need to show that Ag is linear, i.e., that $(Ag)(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 (Ag(f_1)) + \alpha_2 (Ag(f_2))$.

This follows from basic properties of integrals:

$$(Ag)(\alpha_1 f_1 + \alpha_2 f_2) = \int_0^1 g(x)(\alpha_1 f_1(x) + \alpha_2 f_2(x))dx$$

$$= \alpha_1 \int_0^1 g(x) f_1(x)dx + \alpha_2 \int_0^1 g(x) f_2(x)dx = \alpha_1 ((Ag)(f_1)) + \alpha_2 ((Ag)(f_2))$$

B. For any $y \in [0,1]$, there is another map from V to the base field, defined by $(By)(f) = f(y)$. Show that $By \in V^*$*, i.e., that it is a linear map from V* to the base field.

To show $(By)(f+g) = (By)(f) + (By)(g)$: Left hand side is $(f+g)(y)$, which, by definition of addition in V, is $f(y)+g(y)$. Right-hand side is also $f(y)+g(y)$, applying By to f and g separately.

C. For any $y \in [0,1]$, is there a $g \in V$ for which $By = Ag$?

No. We'd need a g for which $(Ag)(f) = (By)(f)$, i.e., $\int_0^1 g(x)f(x)dx = f(y)$. So g would have to be zero for all $x \neq y$. Since V (by definition) only contains smooth functions, then g would have to be zero everywhere,. While we can imagine a generalized function, which has these properties i.e., $\delta(x-y)$, it does not have the smoothness conditions required to be in V.

Q2: Direct path to the trace as an intrinsic property of vector-space homomorphisms (AKA matrices)

This homework is closely modeled after a Quora posting of Senia Sheydvasser. Assistant Professor Department of Mathematics, Bates College

Consider a vector space V and its dual space V^*. Elements in $V^* \otimes V$ are sums of elementary tensor products $\Phi = \phi \otimes v$, for $x \in V$ and $\phi \in V^*$. Thus, elements in $V^* \otimes V$ can also be considered to be in $Hom(V,V)$ (i.e., there is a natural homomorphism L from $V^* \otimes V$ to $Hom(V,V)$). The correspondence L between $V^* \otimes V$ to $Hom(V,V)$ takes an elementary tensor product $\Phi = \phi \otimes v$ to $L(\Phi) \in Hom(V,V)$ given by $L(\Phi)(w) = \phi(w)v$. (Note $\phi(w)$ is a scalar). L is then extended to sums of elementary tensor products by linearity.

A. For $\Phi_1=\phi_1\otimes v_1$ and $\Phi_2=\phi_2\otimes v_2$, determine the action of $L(\Phi_1)\circ L(\Phi_2)$ on an arbitrary $w\in V$, where $\circ$ is composition in $Hom(V,V)$. Express this as the image under L of an elementary tensor product $\Phi_{12}\in V^*\otimes V$. Use this to define a composition rule in $V^*\otimes V$, $\Phi_{12}=\Phi_1\circ\Phi_2$, for which $L(\Phi_{12})=L(\Phi_1)\circ L(\Phi_2)$

We determine, for an arbitrary $w\in V$, the action of $L(\Phi_1)\circ L(\Phi_2)$:
$\left(L(\Phi_1)\circ L(\Phi_2)\right)(w)=L(\Phi_1)\left(L(\Phi_2)(w)\right)$, since $\circ$ is composition.
$L(\Phi_1)\left(L(\Phi_2)(w)\right)=L(\Phi_1)\left(\phi_2(w)(v_2)\right)$, definition of L.
$L(\Phi_1)\left(\phi_2(w)(v_2)\right)=\phi_2(w)L(\Phi_1)(v_2)$, since $L(\Phi_1)$ is linear and $\phi_2(w)$ is a scalar.
$\phi_2(w)L(\Phi_1)(v_2)=\phi_2(w)\phi_1(v_2)v_1$, definition of L.

So $\left(L(\Phi_1)\circ L(\Phi_2)\right)(w)=\phi_2(w)\phi_1(v_2)v_1$. The right-hand side is also $L\left(\left(\phi_1(v_2)\phi_2\right)\otimes v_1\right)(w)$. Since this holds for all w, $\Phi_{12}=\left(\phi_1(v_2)\phi_2\right)\otimes v_1=\phi_1(v_2)\left(\phi_2\otimes v_1\right)$, i.e, another elementary tensor product.

Composition in $V^*\otimes V$ is now defined by $\left(\phi_1\otimes v_1\right)\circ\left(\phi_2\otimes v_2\right)=\phi_1(v_2)\left(\phi_2\otimes v_1\right)$.

B. Determine the $\Phi_{21}\in V^*\otimes V$ for which $L(\Phi_{21})=L(\Phi_2)\circ L(\Phi_1)$.
As in A, $\Phi_{21}=\phi_2(v_1)\left(\phi_1\otimes v_2\right)$
C. There is also a natural mapping T from $V^*\otimes V$ to the base field of scalars, defined by $T(\Phi)=\phi(v)$ for elementary tensor products and extended to all of $V^*\otimes V$ by linearity. Determine $T(\Phi_{12})$ and $T(\Phi_{21})$. What happens?
$T(\Phi_{12})=T\left(\phi_1(v_2)\left(\phi_2\otimes v_1\right)\right)=\phi_1(v_2)T\left(\phi_2\otimes v_1\right)=\phi_1(v_2)\phi_2(v_1)$.
Similarly $T(\Phi_{21})=\phi_2(v_1)\phi_1(v_2)$.
So, $T(\Phi_{12})=T(\Phi_{12})$, i.e., $T(\Phi_1\circ\Phi_2)=T(\Phi_2\circ\Phi_1)$.

D. Now interpret T in coordinates. Specifically, take the "one-hot" basis for V, $v_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},v_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\cdots,v_n=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$, and a typical vector in $x\in V$, $x=\sum_{k=1}^{n}x_kv_k$. Take the one-hot basis for V^*, where ϕ_k maps $\phi_k(v_k)=1$ but $\phi_j(v_k)=0$ for $j\neq k$, and a typical $\varphi\in V^*$, $\varphi=\sum_{k=1}^{n}\varphi_k\phi_k$. Written more compactly, $\phi_j(v_k)=\delta_{j,k}$. Then the $M_{j,k}=\phi_j\otimes v_k$ are a basis for $V^*\otimes V$, and an arbitrary $M\in V^*\otimes V$ can be written as $M=\sum_{j,k=1}^{n}m_{j,k}M_{j,k}$, where the $m_{j,k}$ are the matrix entries for M. Determine $T(M_{j,k})$ and then $T(M)$.

$T(M_{j,k})=T(\phi_j\otimes v_k)=\delta_{j,k}$, by the definition of T in part C.
By linearity of T,

$$T(M) = \sum_{j,k=1}^{n} m_{j,k} M_{j,k} = \sum_{j,k=1}^{n} m_{j,k} T\left(M_{j,k}\right) = \sum_{j,k=1}^{n} m_{j,k} \delta_{j,k} = \sum_{j=1}^{n} m_{j,j} \, .$$

So we've defined the trace in a coordinate-free way, and demonstrated its key property: that the trace of a pairwise composition is independent of the order of composition.

Senia Sheydvasser

PhD in Mathematics · Upvoted by John Brown, Master's in Math, Math instructor and Adam Merberg, Ph.D. Mathematics, University of California, Berkeley · 8y

Originally Answered: Why is tr(AB) = tr(BA) true?

The definition of trace as the sum of the diagonal entries of a matrix is easy to learn and easy to understand. However, it doesn't (a priori) have any nice geometric or other interpretation---it just looks a computation tool. Attacking it from this perspective basically means that you are stuck with computational proofs of facts like $tr(AB) = tr(BA)$.

They aren't *bad*, per se. They are easy to understand, and certainly what should be shown when someone is initially learning linear algebra. There is a deeper reason for why $tr(AB) = tr(BA)$, but it is pretty abstract and in particular requires the tensor product in order to understand.

Consider the space of linear operators from a vectors space V back to itself. If we choose a particular set of coordinates, such operators will look like square matrices. However, we shall aim to avoid coordinates as much as possible.

We denote by V^* the dual space of V, which the space of linear functionals on V---that is, linear maps λ such that if we plug in a vector v, $\lambda(v)$ is a scalar.

If we then take the tensor product $V^* \otimes V$, it is isomorphic to the space of linear operators $V \to V$. The isomorphism works like this: if $w \in V$, then $(\lambda \otimes v)w = \lambda(w)v$.

We can also figure out how composition works out under this isomorphism---recall that composition of linear maps is just the same thing as multiplying the corresponding matrices.

$$\begin{aligned}
&(\lambda_2 \otimes v_2)\left((\lambda_1 \otimes v_1)w\right) \\
&= (\lambda_2 \otimes v_2)\left(\lambda_1(w)v_1\right) \\
&= \lambda_2\left(\lambda_1(w)v_1\right) v_2 \\
&= \lambda_2(v_1)\lambda_1(w)v_2
\end{aligned}$$

hence

$$(\lambda_2 \otimes v_2) \circ (\lambda_1 \otimes v_1) = \lambda_2(v_1)(\lambda_1 \otimes v_2)$$

Now, how does the trace come in? Well, there is a natural map from $V^* \otimes V$ to the field of scalars which works like this: $\lambda \otimes v = \lambda(v)$. The amazing thing is that, if you work everything out in coordinates, this is the trace.

This shows that the trace, far from being some abstract computational tool, is actually a fundamental and natural map in linear algebra. In particular, the above analysis automatically gives a proof that $tr\left(ABA^{-1}\right) = tr(B)$.

But why is the stronger statement $tr(AB) = tr(BA)$ true? Well, let's compute both of them.

$$\begin{aligned}
&tr\left((\lambda_2 \otimes v_2) \circ (\lambda_1 \otimes v_1)\right) \\
&= tr\left(\lambda_2(v_1)(\lambda_1, v_2)\right) \\
&= \lambda_2(v_1)\lambda_1(v_2)
\end{aligned}$$

On the other hand:

$$\begin{aligned}
&tr\left((\lambda_1 \otimes v_1) \circ (\lambda_2 \otimes v_2)\right) \\
&= tr\left(\lambda_1(v_2)(\lambda_2, v_1)\right) \\
&= \lambda_1(v_2)\lambda_2(v_1)
\end{aligned}$$