Math 672 Lecture 13

Why do we care about linear functionals and V'?

1. Linear functionals arise in many contexts:

   • Evaluation (at a point) of continuous functions:

   – E.g. Given $c \in [a,b]$, then

$$e_c : C[a,b] \to \mathbb{R},$$
$$e_c(f) = f(c)$$

   is a linear functional on $C[a,b]$.

   • Integration (or taking the mean):

$$m : C[a,b] \to \mathbb{R},$$
$$m(f) = \int_a^b f(x)\,dx \quad \left(\text{or } m(f) = \frac{1}{b-a}\int_a^b f(x)\,dx\right)$$

   is a linear functional.

   • The trace of a matrix:

$$tr : \mathbb{R}^{n \times n} \to \mathbb{R},$$
$$tr(A) = \sum_{i=1}^{n} A_{ii}$$

   is a linear functional (with interesting properties, e.g. $tr(A)$ = the sum of the eigenvalues of A).

   • Most importantly, derivatives are defined by linear functional.

   Suppose V is a vector space over $\mathbb{R}$ and $f : V \to \mathbb{R}$ is a functional (not necessarily linear). Assuming we have a norm $\|\cdot\|$ defined on V (we discuss norms later in the course), we define the <u>derivative</u> of f at $v \in V$ (if it exists) to be $\ell \in \mathcal{L}(V,\mathbb{R})$ satisfying

$$\lim_{w \to 0} \frac{|f(v+w)-f(v)-\ell(w)|}{\|w\|} = 0.$$

(If such an ℓ does not exist, then f is not differentiable at v.) There are many notations for the derivative of f at v; the most common are probably $f'(v)$ or $Df(v)$. Note that $f'(v) \in \mathcal{L}(V,\mathbb{R})$ (or $Df(v) \in \mathcal{L}(V,\mathbb{R})$).

Question: If $f'(v)$ is a linear functional, then why do we say (in calculus class) that, for $f:\mathbb{R}\to\mathbb{R}$, $f'(x)$ is a <u>number</u>?

Why do we care about the dual map T' of $T \in \mathcal{L}(V,W)$?

This will be more evident when we introduce inner product spaces, in which context the dual map T' becomes the adjoint T^*.

However, here is one important application: Suppose $T \in \mathcal{L}(V,W)$ and $f: W \to \mathbb{R}$ (f is presumably nonlinear). Then $g = f \circ T$ maps V into $\mathbb{R}$, and we can ask for $g'(v) \in V'$. By the chain rule,

$$g'(v) = f'(T(v)) \circ T = T'(f'(T(v)))$$

$(f'(T(v)) \in W'$ and $T' : W' \to V'$, so $T'(f'(T(v))) \in V'$, as expected$)$.

I wish to prove two facts about dual maps.

<u>Theorem</u>: $\dim(R(T')) = \dim(R(T))$.

<u>Proof</u>: Choose $v_1, \ldots, v_n \in V$ such that $\{T(v_1), \ldots, T(v_n)\}$ is a basis for $R(T)$. Extend it to a basis $\{w_1, \ldots, w_m\}$ for W, where $w_j = T(v_j)$ for $j = 1, 2, \ldots, n \leq m$. Let $\{\psi_1, \ldots, \psi_m\}$ be the dual basis for W'. We claim that $\{T'(\psi_1), \ldots, T'(\psi_n)\}$ is a basis for $R(T')$. First, $\{T'(\psi_1), \ldots, T'(\psi_n)\}$ is linearly independent:

$$\alpha_1 T'(\psi_1) + \cdots + \alpha_n T'(\psi_n) = 0$$

$$\Rightarrow T'(\alpha_1 \psi_1 + \cdots + \alpha_n \psi_n) = 0$$

$$\Rightarrow ((\alpha_1 \psi_1 + \cdots + \alpha_n \psi_n) \circ T)(v) = 0 \quad \forall\, v \in V$$

$$\Rightarrow (\alpha_1 \psi_1 + \cdots + \alpha_n \psi_n)(T(v_j)) = 0 \quad \forall\, j = 1, \ldots, n$$

$$\Rightarrow \sum_{i=1}^{n} \alpha_i \psi_i(T(v_j)) = 0 \quad \forall\, j = 1, \ldots, n$$

$$\Rightarrow \alpha_j = 0 \quad \forall\, j = 1, \ldots, n.$$

Now suppose $\varphi \in \mathcal{R}(T')$, say $\varphi = T'(\psi)$ for some $\psi \in W'$.

We have

$$\varphi = \psi \circ T$$

$$\Rightarrow \varphi = (\alpha_1 \psi_1 + \cdots + \alpha_m \psi_m) \circ T \text{ for some } \alpha_1, \ldots, \alpha_m$$

$$\Rightarrow \varphi(v) = \sum_{j=1}^{m} \alpha_j \psi_j (T(v)) \quad \forall v \in V$$

Fix $v \in V$; then $T(v) \in \mathcal{R}(T)$, so $T(v) = \sum_{i=1}^{n} \beta_i T(v_i)$

for some $\beta_1, \ldots, \beta_n$. It follows that

$$\psi_j(T(v)) = 0 \quad \forall j = n+1, \ldots, m$$

$$\Rightarrow \varphi(v) = \sum_{j=1}^{n} \alpha_j \psi_j (T(v))$$

$$= ((\alpha_1 \psi_1 + \cdots + \alpha_n \psi_n) \circ T)(v)$$

Thus

$$\varphi = (\alpha_1 \psi_1 + \cdots + \alpha_n \psi_n) \circ T$$

$$= \alpha_1 \psi_1 \circ T + \cdots + \alpha_n \psi_n \circ T$$

$$= \alpha_1 T'(\psi_1) + \cdots + \alpha_n T'(\psi_n)$$

$$\Rightarrow \varphi \in \operatorname{span}(T'(\psi_1), \ldots, T'(\psi_n)). \quad //$$

Theorem Let V, W be finite-dimensional vector spaces over a field F with bases $\mathcal{B}, \mathcal{C}$, respectively, and let $\mathcal{B}', \mathcal{C}'$ be the dual bases of $\mathcal{B}, \mathcal{C}$, respectively. Then

$$\mathcal{M}_{\mathcal{C}',\mathcal{B}'}(T') = \mathcal{M}_{\mathcal{B},\mathcal{C}}(T)^T \qquad \text{(matrix transpose).}$$

Proof: Let us write $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$, $\mathcal{C} = \{w_1, w_2, \ldots, w_m\}$, $\mathcal{B}' = \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$, $\mathcal{C}' = \{\psi_1, \psi_2, \ldots, \psi_m\}$, $A = \mathcal{M}_{\mathcal{B},\mathcal{C}}(T)$, $B = \mathcal{M}_{\mathcal{C}',\mathcal{B}'}(T')$.

Let $v \in V$, $\psi \in W'$ be given, and note that

$$\psi(T(v)) = (T'(\psi))(v).$$

Suppose

$$v = \sum_{j=1}^{n} x_j v_j, \quad \psi = \sum_{i=1}^{m} y_i \psi_i.$$

Then

$$T(v) = \sum_{k=1}^{m} (Ax)_k w_k \quad (\text{since } \mathcal{M}_{\mathcal{C}}(T(v)) = A\mathcal{M}_{\mathcal{B}}(v))$$

and

$$\begin{aligned}
\psi_i(T(v)) &= \psi_i\left(\sum_{k=1}^{m} (Ax)_k w_k\right) \\
&= \sum_{k=1}^{m} (Ax)_k \psi_i(w_k) \\
&= (Ax)_i \quad (\text{since } \psi_i(w_k) = \delta_{ik})
\end{aligned}$$

$$\Rightarrow \psi(T(v)) = \left(\sum_{i=1}^{m} y_i \psi_i\right)(T(v))$$

$$= \sum_{i=1}^{m} y_i \psi_i(T(v))$$

$$= \sum_{i=1}^{m} y_i (Ax)_i$$

$$= \sum_{i=1}^{n} y_i \left(\sum_{j=1}^{n} A_{ij} x_j\right)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} x_j y_i .$$

On the other hand,

$$T'(\psi) = \sum_{j=1}^{n} (By)_j \varphi_j \quad (\text{since } \mathcal{M}_{\mathcal{B}'}(T'(\psi)) = B\mathcal{M}_{\mathcal{C}'}(\psi))$$

$$\Rightarrow (T'(\psi))(v) = \left(\sum_{j=1}^{n} (By)_j \varphi_j\right)(v)$$

$$= \sum_{j=1}^{n} (By)_j \varphi_j(v).$$

Since

$$\varphi_j(v) = \varphi_j\left(\sum_{k=1}^{n} x_k v_k\right)$$

$$= \sum_{k=1}^{n} x_k \varphi_j(v_k)$$

$$= x_j \quad (\text{since } \varphi_j(v_k) = \delta_{jk}),$$

we obtain

$$(T'(\psi))(v) = \sum_{j=1}^{n} (By)_j x_j$$

$$= \sum_{j=1}^{n} \left(\sum_{i=1}^{m} B_{ji} y_i \right) x_j$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{m} B_{ji} x_j y_i$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} B_{ji} x_j y_i .$$

Thus

$$\psi(T(v)) = (T'(\psi))(v)$$

$$\Rightarrow \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} x_j y_i = \sum_{i=1}^{m} \sum_{j=1}^{n} B_{ji} x_j y_i$$

Now, this holds for all $v \in V$ and all $\psi \in W'$, that is, for all $x \in F^n$ and all $y \in F^m$. Therefore,

$$B_{ji} = A_{ij} \quad \forall\, i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n,$$

that is, $B = A^T$. //