Notes on the Dual Space

Let $V$ be a vector space over a field $F$. The **dual space** of $V$ is the vector space $V^* = \mathrm{Lin}(V, F)$ consisting of the linear mappings $\phi : V \to F$. The elements of $V^*$ are called linear forms or linear functionals. If $V$ has basis $e = (e_i)_{i\in I}$ then any vector $v \in V$ can be written uniquely in the form $v = \sum_{i\in I} x_i e_i$ with $x_i = 0$ for all but a finite number of $i \in I$. If $\phi \in V^*$ then $\phi(v) = \sum_{i\in I} a_i x_i$ where $a_i = f(e_i)$. Conversely, if $(a_i)_{i\in I}$ is a family of scalars indexed by I then

$$\phi(v) = \sum_{i\in I} a_i x_i$$

defines a linear form $\phi$ on $V$ with $\phi(e_i) = a_i$. In particular, $e_i^*(v) = x_i$ defines a linear form on $V$ called the $i$-th coordinate function for the basis $e = (e_i)_{i\in I}$. If $I$ is a finite set then any $\phi \in V^*$ can be uniquely written in the form

$$\phi = \sum_{i\in I} a_i e_i^*$$

with $a_i \in F$. We have $a_i = \phi(e_i)$ since $e_i^*(e_j) = \delta_{ij}$ (the Kronecker delta). Hence, in this case, $e^* = (e_i^*)_{i\in I}$ is a basis for $V^*$, the basis of $V^*$ dual to the basis $e = (e_i)_{i\in I}$ or **dual basis**. We thus have the result

**Theorem 1.** *If $V$ is a finite-dimensional vector space then* $\dim(V) = \dim(V^*)$.

If the basis $(e_i)$ is infinite then the subspace $W$ of $V^*$ spanned by the vectors $e_i^*(i \in I)$ is a proper subspace of $V$. For example, the linear form $\phi$ with $\phi(e_i) = 1$ for all $i$ is not in $W$ since $\phi(e_i) = 0$ for all but a finite number of $i$ if $\phi \in W$. In fact, using set theory, one can show that $\dim(V^*) > \dim(V)$ if $\dim(V)$ is not finite.

**Example 1.** *If $V = F^X$ is the set of $F$-valued functions on $X$ then, for any $x \in X$, the function $e_x :\to F$ defined by $e_x(f) = f(x)$ (evaluation at $x$) is a linear form on $V$.*

**Example 2.** *If $V = C_{\mathbb{R}}([a, b])$ is the vector space of continuous real-valued functions on the interval $[a, b]$ then*

$$\phi(f) = \int_a^b f(x)\,dx$$

*defines a linear form on $V$.*

**Example 3.** *If $V = F^n$ and $\phi \in V^*$ then*

$$\phi(x_1, \dots, x_n) = a_1x_1 + \cdots a_nx_n = [a_1, a_2, \dots, a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

*where $[a_1, a_2, \dots, a_n] = [\phi]$ is the matrix of $\phi$.*

If $S$ is a subset of the vector space $V$ then the **annihilator** of $S$ is the set

$$S^0 = \{\phi \in V^* \mid \phi(s) = 0 \text{ for all } s \in S\}$$

consisting of those linear forms on $V$ which vanish on $S$. We have

$$S^0 = (\mathrm{span}(S))^0.$$

If $W$ is a subspace of $F^n$ then the elements of $W^0$ give the hyperplanes $a_1x_1 + \cdots + a_nx_n = 0$ which contain $W$.

**Example 4.** *If* $W = \mathrm{Span}((1,1,1))$ *then* $W^0 = \mathrm{Span}(\phi_1, \phi_2)$ *where* $\phi_1(x_1, x_2, x_3) = x_1 - x_2$, $\phi_2(x_1, x_2, x_3) = x_2 - x_3$.

If $W$ is a subspace of $V$ we can define a linear mapping $T : W^0 \to (V/W)^*$ by

$$T(\phi)(v + W) = \phi(v).$$

The function $T(\phi)$ is well defined on $V/W$ since $\phi$ vanishes on $W$. That $T(\phi)$ is linear is left to the reader as well as the fact that $T$ is injective. If $\psi$ is a linear form on $V/W$ then $\phi(v) = \psi(v + W)$ defines a linear form on $V$ and $\psi = T(\phi)$ which shows that $T$ is surjective and hence an isomorphism.

**Theorem 2.** *If $W$ is a subspace of the vector space $V$ then*

$$W^0 \cong (V/W)^*.$$

*If* $\dim(V/W) < \infty$ *then*

$$\dim(W^0) = \mathrm{codim}(W),$$

*where* $\mathrm{codim}(W) = \dim(V/W)$. *If* $\dim(V) < \infty$, *we have*

$$\dim(W^0) = \dim(V) - \dim(W).$$

There is a canonical mapping $R$ of a vector space $V$ into its second dual $V^{**} = (V^*)^*$ defined by $R(v) = v^{**}$ where $v^{**}(\phi) = \phi(v)$. The proof of the linearity of $v^{**}$ and $R$ are left to the reader. If $R(v) = 0$ we have $\phi(v) = 0$ for all $\phi \in V^*$. If $v \neq 0$ then it can be completed to a basis $B$ of $V$. Then the linear form $v^*$ which is 1 on $v$ and 0 on the other vectors in $B$ contradicts the fact that $\phi(v) = 0$ for all $\phi \in V^*$. If $V$ is finite-dimensional, the mapping $R$ is surjective since $\dim(V) = \dim(V^{**})$. In this case, we use the isomorphism $R$ to identify $V$ with its second dual.

If $W$ is a subspace of $V^*$, we define the annihilator of $W$ in $V$ to be the set

$$W^0 = \{v \in V : \phi(v) = 0 \text{ for all } \phi \in W\}.$$

If $\dim(V) < \infty$ then $R(W^0) = \{v^{**} | v^{**}(\phi) = 0 \text{ for all } \phi \in W\}$ is the annihilator of $W$ in $V^{**}$ which gives the following result.

**Theorem 3.** *If $V$ is finite-dimensional and $W$ is a subspace of $V^*$ then* $\dim(W^0) = \mathrm{codim(W)}$.

This can also be proven by showing that the mapping $v \mapsto v_*$ of $W^0$ into $(V^*/W)^*$, where $v_*(\phi + W) = \phi(v)$ is an isomorphism when $V$ is finite-dimensional.

**Corollary 4.** *If* $\dim(V) < \infty$ *and $W$ is a subspace of $V$ or $V^*$ then* $W^{00} = W$.

This follows from the fact that $W \subseteq W^{00}$. Since $W_1 \subseteq W_2$ implies $W_2^0 \subseteq W_1^0$ we get, in the case $V$ is finite-dimensional, an inclusion reversing bijection $W \mapsto W^0$ between the the subspaces of $V$ and the subspaces of $W$. The corresponding subspaces are said to be dual to one another.

**Theorem 5.** *If $W_1, W_2$ are subspaces of a vector space $V$ we have* $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$. *If* $\dim(V) < \infty$ *we also have* $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.

The proof is left to the reader.

If $T : U \to V$ is a linear mapping the transpose of $T$ is the mapping $T^t : V^* \to U^*$ defined by

$$T^t(\phi) = \phi \circ T.$$

The proof that $T^t$ is linear is left to the reader as is the fact that the mapping $T \mapsto T^t$ is linear. If $U, V$ are finite-dimensional then $(T^t)^t(u^{**}) = u^{**} \circ T^t$ and

$$u^{**} \circ T^t(\phi) = u^{**} \circ \phi \circ T = \phi(T(u) = T(u)^{**}(\phi)$$

for $\phi \in V^{**}$ so that $(T^t)^t = T$ after identifying $U$ and $V$ with their duals. If $S : V \to W$ then $(ST)^t = T^t S^t$ since

$$(ST)^t(\phi) = \phi \circ ST = S^t(\phi) \circ T = T^t(S^t(\phi)) = T^t S^t(\phi)$$

for $\phi \in U^*$.

**Theorem 6.** *If $T : U \to V$ is linear then*

$$(\mathrm{Im(T)})^0 = \mathrm{Ker(T^t)} \text{ and } (\mathrm{Im(T^t)})^0 = \mathrm{Ker(T)}.$$

The proof is left to the reader.

If $U, V$ are vector spaces with bases $e = (e_1, \ldots, e_)$, $f = (f_1, \ldots, f_m)$ respectively and $T : U \to V$ is a linear mapping then

$$T(e_j) = \sum_{i=1}^{m} a_{ij} f_i.$$

The $m \times n$ matrix $A = [a_{ij}]$ is the matrix $[T]_e^f$ of $T$ with respect to the bases $e, f$. If $e^* = (e_1^*, \ldots, e_n^*)$, $f^* = (f_1^*, \ldots, f_m^*)$ are the corresponding dual bases then

$$T^t(f_i^*) = \sum_{j=1}^{n} b_{ji} e_j^*.$$

The matrix $n \times m$ matrix $B = [b_{ji}]$ is the matrix $[T^t]_{f^*}^{e^*}$ of $T^t$ with respect to the bases $f^*, e^*$. Since

$$b_{ji} = T^t(f_i^*)(e_j) = f_i^*(T(e_j)) = a_{ij},$$

we see the the $(j, i)$-th entry of $B$ is the $(i, j)$-th entry of $A$. The matrix $B$ is called the transpose of $A$ and is denoted by $A^t$.

**Theorem 7.** *If $U, V$ are vector spaces with bases $e = (e_1, \ldots, e_)$, $f = (f_1, \ldots, f_m)$ respectively and $T : U \to V$ is a linear mapping then*

$$[T^t]_{f^*}^{e^*} = ([T]_e^f)^t.$$

The linearity of the transpose for matrices as well as $(A^t)^t = A$ and $(AB)^t = B^t A^t$ follow from the above.

If $A \in F^{m \times n}$, we have two linear mappings $T_A : F^{n \times 1} \to F^{m \times 1}$, $T_A : F^{1 \times m} \to F^{1 \times n}$ defined by

$$T_A(X) = AX, \quad T^A(Y) = AY.$$

If we identify $(F^{n \times 1})^*$ with $F^{1 \times n}$ using the isomorphism $\phi \mapsto [\phi]$ then $(T_A)^t = T^A$. We have

$$\mathrm{Ker(T_A) = Null(A), \quad Im(T_A) = Col(A), \quad Ker(T^A) \cong Null(A^t), \quad Im(T^A) = Row(A) \cong Col(A^t)}$$

where the isomorphisms are defined by the transpose. If $W$ is a subset of $F^{n \times 1}$ then

$$W^0 = \{Y \in F^{1 \times n} \mid YX = 0 \text{ for all } X \in W\}$$

and if $W$ is a subset of $F^{1 \times n}$ then

$$W^0 = \{X \in F^{n \times 1} \mid YX = 0 \text{ for all } Y \in W\}.$$

Since $\mathrm{(Im(T_A)^0 = Null(T^A)}$ and $\mathrm{(Im(T^A))^0 = Ker(T_A)}$, we have

$$\mathrm{Col(A)^0 = \{Y \in F^{1 \times m} \mid YA = 0\} = \{Y \in F^{1 \times m} \mid A^t Y^t = 0\}, \quad Row(A)^0 = Null(A)}$$

It follows that

$$\mathrm{rank(A) + nullity(A^t) = m, \quad rank(A^t) + nullity(A) = n.}$$

In particular, $\mathrm{rank(A) = m - nullity(A^t) = rank(A^t)}$. Since $\mathrm{Row(A) = Row(A)^{00}}$, we also have

$$\mathrm{Row(A) = Null(A)^0}.$$

Thus two $m \times n$ matrices which have the same also have the same row space.

There is a canonical form for the row space of a matrix $A$, namely, its **reduced echelon form** REF($A$). This $m \times n$ matrix is row equivalent to $A$ and so has the same row space as $A$. It has the following properties:

(a) The non-zero rows precede the zero rows;

(b) The first non-zero entry in a non-zero row is 1;

(c) If the first-nonzero entry in row $i$ occurs in column $j_i$ then $j_1 < j_2 < \cdots < j_s$, where $s$ is the number of non-zero rows;

(d) For $1 \leq i \leq s$, there is exactly one non-zero entry in column $j_i$.

The natural number $s$ is the dimension of the row space of $A$.

**Theorem 8.** *Let $A, A'$ be $m \times n$ matrices over a field $F$. Then the following are equivalent:*

(1) *There an invertible matrix $P$ such that $A' = PA$;*

(2) $\mathrm{Row(A)} = \mathrm{Row(A')}$*;*

(3) $\mathrm{REF}(A) = \mathrm{REF}(A')$.

*Proof.* As (1) $\implies$ (2) and (3) $\implies$ (1) are immediate, we only give the proof of (2) $\implies$ (3). We proceed by induction on the common number of non-zero rows of $A$ and $B$. Let $A_1, \ldots, A_s$ be the non-zero rows of $A$ and let $A'_1, \ldots, A'_s$ be the non-zero rows of $A'$. Let $j_1, \ldots, j_s$ and $j'_1, \ldots, j'_s$ be the distinguished columns of $A$ and $B$ respectively. Since $A_1 = A'_1 + a_2 A'_2 + \ldots + A'_s$ we have $j_1 = j'_1$. It follows that

$$\mathrm{Span(A_2, \ldots, A_s) = Span(A'_2, \ldots, A'_s)}.$$

Thus the matrices obtained from $A$ and $A'$ by deleting their first row have the same row space and are in reduced echelon form. By our inductive hypothesis, these two matrices are equal. It follows that $A_1 = A'_1$. $\square$

(last modified 2:45pm March 10, 2004)