The Dual Space

Hopefully the reader has seen the idea of the *dual space* before, otherwise this section will probably be a little confusing.

Let V be a finite dimensional vector space over the real numbers $\mathbb{R}$ A **linear functional** is a function $f : V \to \mathbb{R}$ such that for all vectors $\mathbf{u}, \mathbf{v} \in V$ and scalars $\alpha, \beta \in \mathbb{R}$,

$$f(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha f(\mathbf{u}) + \beta f(\mathbf{v}).$$

The **dual space** of V, denoted V^*, is the set of all **linear functionals** on V. That is,

$$V^* = \{f : V \to \mathbb{R} \mid f \text{ is linear}\}.$$

The dual space V^* is itself a vector space over $\mathbb{R}$, with vector addition and scalar multiplication defined *pointwise*:

$$(f + g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v}),$$
$$(\alpha f)(\mathbf{v}) = \alpha\, f(\mathbf{v}),$$

for all $f, g \in V^*$, $\alpha \in \mathbb{R}$, and $\mathbf{v} \in V$.

Key Point

One of the most important properties of the dual space of V is that the standard dot product allows us to associate every vector $\mathbf{v} \in V$ with a linear functional $f_{\mathbf{v}} \in V^*$ via the rule

$$\boxed{f_{\mathbf{v}}(\mathbf{y}) = \mathbf{v} \cdot \mathbf{y} = \mathbf{v}^\top \mathbf{y}, \quad \text{for all } \mathbf{y} \in V}.$$

Theorem 1 (Finite-Dimensional Riesz Representation Theorem). *Let $V = \mathbb{R}^n$ with the standard inner product. Then the map*

$$\psi : V \to V^*, \quad \psi(\mathbf{v})(\mathbf{y}) = \mathbf{v}^\top \mathbf{y},$$

is a vector space isomorphism.

Proof. First, linearity: for all $\alpha, \beta \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w}, \mathbf{y} \in V$,

$$f_{\alpha\mathbf{v}+\beta\mathbf{w}}(\mathbf{y}) = (\alpha\mathbf{v} + \beta\mathbf{w})^\top \mathbf{y} = \alpha\mathbf{v}^\top \mathbf{y} + \beta\mathbf{w}^\top \mathbf{y} = \alpha f_{\mathbf{v}}(\mathbf{y}) + \beta f_{\mathbf{w}}(\mathbf{y}).$$

Next, injectivity: if $f_{\mathbf{v}} = 0$, then $\mathbf{v}^\top \mathbf{y} = 0$ for all $\mathbf{y}$. In particular, $\mathbf{v}^\top \mathbf{v} = \|\mathbf{v}\|^2 = 0$, so $\mathbf{v} = \mathbf{0}$.

Finally, surjectivity: given $g \in V^*$, define $v_i = g(\mathbf{e}_i)$ for the standard basis $\{\mathbf{e}_i\}$. Set $\mathbf{v} = \sum_i v_i \mathbf{e}_i$. Then for each basis vector,

$$f_{\mathbf{v}}(\mathbf{e}_i) = \mathbf{v}^\top \mathbf{e}_i = v_i = g(\mathbf{e}_i).$$

Since $f_{\mathbf{v}}$ and g agree on a basis, they are equal.

Thus ψ is linear, injective, and surjective, hence an isomorphism. □

Corollary 1. $\dim V^* = \dim V$ *for finite-dimensional* V

As a consequence of the above, suppose that $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis for V, then there is a uniquely associated **dual basis** $\{\mathbf{e}^1, \ldots, \mathbf{e}^n\} \subseteq V^*$ defined by

$$\mathbf{e}^i(\mathbf{e}_j) = \delta^i_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Every linear functional $f \in V^*$ can be uniquely expressed as

$$f = \sum_{i=1}^{n} f(\mathbf{e}_i)\, \mathbf{e}^i.$$

Some intuition

It can sometimes help to think of vectors in V as column vectors and vectors in V^* as row vectors. The dot product is now the application of a linear functional to a vector. The transpose is the isomoprhism between V and V^*.

The Adjoint

The reason why we have gone to all the trouble of introducing the dual space is that the transpose of a matrix is the natural matrix representation of the **adjoint** of a linear map.

The Adjoint of a Linear Map

Let $T : V \to W$ be a linear map between finite-dimensional real vector spaces. The **dual map** $T^* : W^* \to V^*$ is defined by

$$(T^* f)(\mathbf{v}) = f(T\mathbf{v}) \quad \text{for all } f \in W^*,\ \mathbf{v} \in V.$$

In words: to evaluate $T^* f$ at a vector $\mathbf{v}$, first apply T to $\mathbf{v}$, then apply the functional f to the result.

Theorem 2. *The matrix of the dual map T^* is the transpose of the matrix of T.*

Proof. Choose bases:

- $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for V,
- $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ for W,
- $\mathcal{B}^* = \{\mathbf{v}^1, \ldots, \mathbf{v}^n\}$ and $\mathcal{C}^* = \{\mathbf{w}^1, \ldots, \mathbf{w}^m\}$ for the dual bases.

Suppose the matrix of T with respect to $\mathcal{B}, \mathcal{C}$ is $A = (a_{ij})$, then the image of v_1 is the first column of A:

$$T(\mathbf{v}_1) = \sum_{i=1}^{m} A_{i1}\, \mathbf{w}_i$$

More generally, the image of v_j is the jth column of A:

$$T(\mathbf{v}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{w}_i.$$

We compute the matrix of T^* with respect to $\mathcal{C}^*, \mathcal{B}^*$. For any $k \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$,

$$(T^*\mathbf{w}^k)(\mathbf{v}_j) = \mathbf{w}^k(T\mathbf{v}_j) = \mathbf{w}^k\left(\sum_{i=1}^{m} a_{ij}\mathbf{w}_i\right) = a_{kj}.$$

On the other hand, if the matrix of T^* is $B = (b_{\ell k})$, then

$$T^*(\mathbf{w}^1) = \sum_{\ell=1}^{n} b_{\ell 1}\,\mathbf{v}^\ell$$

More generally:

$$T^*(\mathbf{w}^k) = \sum_{\ell=1}^{n} b_{\ell k}\mathbf{v}^\ell,$$

so

$$(T^*\mathbf{w}^k)(\mathbf{v}_j) = b_{jk}.$$

Comparing both expressions gives $b_{jk} = a_{kj}$, so $B = A^\top$.

That is the matrix of the dual map T^* is the transpose of the matrix of T. $\square$

The Four Fundamental Subspaces

Recall from Chapter 1 that the key orthogonal decompositions are:

$$\mathbb{R}^n = \operatorname{Row}(A) \oplus \operatorname{Nul}(A), \quad \mathbb{R}^m = \operatorname{Col}(A) \oplus \operatorname{Nul}(A^\top),$$

or equivalently:

$$\operatorname{Nul}(A) = \operatorname{Row}(A)^\perp, \quad \operatorname{Nul}(A^\top) = \operatorname{Col}(A)^\perp.$$

The Annihilator

Let V be a finite-dimensional vector space, and let $W \subseteq V$ be a subspace. The **annihilator** of W is:

$$W^\circ = \{f \in V^* : f(\mathbf{w}) = 0 \text{ for all } \mathbf{w} \in W\},$$

where V^* is the **dual space** (the space of linear functionals on V).

Some notation

If one is not careful then it is easy to confuse the vector space V with the vector spacd V^*. For this reason we will use the following notation. Suppose that $A \subseteq V$ and $B \subseteq V^* R$ then we write:

$$A \cong B$$

When we mean to say that:

$$A = \psi^{-1}[B]$$

where ψ is the isomorphism defined in Theorem 1. We come now to the most important result in this Chapter.

Theorem 3.

$$\mathrm{Nul}(A^\top) \cong (\mathrm{Col}(A))^\circ, \qquad \mathrm{Nul}(A) \cong (\mathrm{Row}(A))^\circ,$$

where the annihilator is taken inside the appropriate dual space, and the isomorphism is the one induced by the dot product.

Proof. We only need to prove one of these results, and the second one follows by symmetry.

Let A be an $m \times n$ real matrix. We will prove that, under the natural identification of vectors in $\mathbb{R}^m$ with linear functionals on $\mathbb{R}^m$ provided by the dot product, the null space of $A^\top$ corresponds, under isomorphism, to the annihilator of the column space of A. That is,

$$\mathrm{Nul}(A^\top) \cong (\mathrm{Col}(A))^\circ.$$

We proceed step by step, explaining the meaning of each concept and how they relate.

The **column space** of A is the subspace of $\mathbb{R}^m$ defined by

$$\mathrm{Col}(A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m.$$

Its **annihilator** is the set of all linear functionals on $\mathbb{R}^m$ that vanish on every vector in $\mathrm{Col}(A)$:

$$(\mathrm{Col}(A))^\circ = \{f \in (\mathbb{R}^m)^* \mid f(\mathbf{y}) = 0 \text{ for all } \mathbf{y} \in \mathrm{Col}(A)\}.$$

Because every $\mathbf{y} \in \mathrm{Col}(A)$ can be written as $\mathbf{y} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$, we can rephrase this condition as:

$$f \in (\mathrm{Col}(A))^\circ \quad \iff \quad f(A\mathbf{x}) = 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n.$$

Now consider a vector $\mathbf{v} \in \mathbb{R}^m$, and let $f_\mathbf{v}$ be the corresponding functional: $f_\mathbf{v}(\mathbf{y}) = \mathbf{v}^\top\mathbf{y}$. We ask: *When does $f_\mathbf{v}$ belong to $(\mathrm{Col}(A))^\circ$?*

By the characterization above, this happens precisely when

$$f_\mathbf{v}(A\mathbf{x}) = 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

Substituting the definition of $f_\mathbf{v}$, this becomes:

$$\mathbf{v}^\top(A\mathbf{x}) = 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^n. \tag{1}$$

Step 4: Rewrite the condition using properties of the transpose

Using the fact that $(XY)^\top = Y^\top X^\top$, we may rewrite this as:

$$\mathbf{v}^\top A\mathbf{x} = (A^\top\mathbf{v})^\top\mathbf{x}.$$

Thus, condition (1) is equivalent to:

$$(A^\top\mathbf{v})^\top\mathbf{x} = 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

In order for this to be true for all x the linear functional $f_{A^T\mathbf{v}}$ must be zero, which iimplies that $A^T v = 0$ which implies that $v \in \mathrm{Nul}(A^T)$.

We have shown the following chain of equivalences:

$$\begin{aligned}
\mathbf{v} \in \mathrm{Nul}(A^\top) &\iff A^\top \mathbf{v} = \mathbf{0} \\
&\iff (A^\top \mathbf{v})^\top \mathbf{x} = 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n \\
&\iff \mathbf{v}^\top A\mathbf{x} = 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n \\
&\iff f_{\mathbf{v}}(A\mathbf{x}) = 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n \\
&\iff f_{\mathbf{v}} \in (\mathrm{Col}(A))^\circ.
\end{aligned}$$

Therefore, under the isomorphism $\mathbf{v} \leftrightarrow f_{\mathbf{v}}$, the subspace $\mathrm{Nul}(A^\top) \subseteq \mathbb{R}^m$ corresponds, under isomorphism, to the subspace $(\mathrm{Col}(A))^\circ \subseteq (\mathbb{R}^m)^*$.

□

Geometric Interpretation

This result has a clean geometric meaning: a vector $\mathbf{v} \in \mathbb{R}^m$ is orthogonal (with respect to the dot product) to every vector in the column space of A if and only if $A^\top \mathbf{v} = \mathbf{0}$. But "orthogonal to the column space" is precisely what it means for the functional $f_{\mathbf{v}}(\mathbf{y}) = \mathbf{v} \cdot \mathbf{y}$ to vanish on $\mathrm{Col}(A)$ — i.e., to be in the annihilator. Thus, in $\mathbb{R}^m$, the annihilator $(\mathrm{Col}(A))^\circ$ is naturally identified with the orthogonal complement $\mathrm{Col}(A)^\perp$, and we recover the familiar fundamental theorem of linear algebra:

$$\mathrm{Nul}(A^\top) = \mathrm{Col}(A)^\perp.$$

Theorem 4. $\mathrm{Nul}(A) \cong (\mathrm{Row}(A))^\circ$

Proof. This follows immediately by symmetry. □

Conclusion. Via the isomorphism ψ between $\mathbb{R}^k$ $(\mathbb{R}^k)^*$, we have the following identities:

$$\boxed{\mathrm{Nul}(A^\top) \cong (\mathrm{Col}(A))^\circ}, \qquad \boxed{\mathrm{Nul}(A) \cong (\mathrm{Row}(A))^\circ}.$$

The Annihilator reveals that the four subspaces reflect a fundamental duality. The matrix $A : \mathbb{R}^n \to \mathbb{R}^m$ induces a **dual map** $A^* : (\mathbb{R}^m)^* \to (\mathbb{R}^n)^*$, and we have:

$$\ker(A^*) = (\operatorname{Im} A)^\circ, \qquad \operatorname{Im}(A^*) = (\ker A)^\circ.$$

When we identify $(\mathbb{R}^k)^* \cong \mathbb{R}^k$ via the isomorphism ψ, these become:

$$\operatorname{Nul}(A^\top) = \operatorname{Col}(A)^\perp, \qquad \operatorname{Row}(A) = \operatorname{Nul}(A)^\perp.$$

Hopefully this chapter has been illuminating.