The Four Fundamental Subspaces and Their Orthogonality Relations

The goal of this section

The goal of this section is to understand the *four fundamental subspaces* associated with a finite dimensional matrix A over the real numbers $\mathbb{R}$. We begin by reviewing matrix multiplication, then define our four subspaces and prove our main theorem. Next we introduce the dual space and adjoint operators. We define the annihilator of a subspace and give an alternative proof of the main theorem. Finally we prove the rank-nullity theorem and also prove that the rank of the transpose of a matrix is equal to the rank of the original matrix.

Matrix Multiplication

Let us recall briefly the formula for the multiplication of a matrix by a vector. If $Bv = w$ then $w_k = \sum_j B_{kj} v_j$. This can be interpreted in two different ways. Firstly as the dot product of each of the rows of B with the vector v.

$$\begin{bmatrix} - & r_1 & - \\ - & r_2 & - \\ - & r_3 & - \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} r_1.v \\ r_2.v \\ r_3.v \end{bmatrix}$$

Secondly as a linear combination of the columns.

$$\begin{bmatrix} | & | & | \\ c_1 & c_2 & c_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = v_1 \begin{bmatrix} | \\ c_1 \\ | \end{bmatrix} + v_2 \begin{bmatrix} | \\ c_2 \\ | \end{bmatrix} + v_3 \begin{bmatrix} | \\ c_3 \\ | \end{bmatrix}$$

It is worth pausing here for a moment to make sure you are clear on this. Here is an example:

Example

Let

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}.$$

Denote the rows of B by

$$r_1 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}, \quad r_2 = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}, \quad r_3 = \begin{bmatrix} 7 & 8 & 9 \end{bmatrix},$$

and the columns by

$$c_1 = \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \quad c_2 = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}, \quad c_3 = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}.$$

1. Row View: Dot Products

Each entry of $\mathbf{w} = B\mathbf{v}$ is the dot product of a row of B with $\mathbf{v}$:

$$\begin{aligned} w_1 &= r_1 \cdot \mathbf{v} = (1)(2) + (2)(-1) + (3)(1) = 2 - 2 + 3 = 3, \\ w_2 &= r_2 \cdot \mathbf{v} = (4)(2) + (5)(-1) + (6)(1) = 8 - 5 + 6 = 9, \\ w_3 &= r_3 \cdot \mathbf{v} = (7)(2) + (8)(-1) + (9)(1) = 14 - 8 + 9 = 15. \end{aligned}$$

Thus,

$$\begin{bmatrix} - & r_1 & - \\ - & r_2 & - \\ - & r_3 & - \end{bmatrix} \mathbf{v} = \begin{bmatrix} r_1 \cdot \mathbf{v} \\ r_2 \cdot \mathbf{v} \\ r_3 \cdot \mathbf{v} \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 15 \end{bmatrix}.$$

2. Column View: Linear Combination

The product $B\mathbf{v}$ is a linear combination of the columns of B, weighted by the entries of $\mathbf{v}$:

$$B\mathbf{v} = v_1 c_1 + v_2 c_2 + v_3 c_3 = 2c_1 + (-1)c_2 + 1c_3.$$

Compute:

$$2c_1 = \begin{bmatrix} 2 \\ 8 \\ 14 \end{bmatrix}, \quad -1c_2 = \begin{bmatrix} -2 \\ -5 \\ -8 \end{bmatrix}, \quad 1c_3 = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}.$$

Add them:

$$\begin{bmatrix} 2 \\ 8 \\ 14 \end{bmatrix} + \begin{bmatrix} -2 \\ -5 \\ -8 \end{bmatrix} + \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 2-2+3 \\ 8-5+6 \\ 14-8+9 \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 15 \end{bmatrix}.$$

So,

$$\begin{bmatrix} | & | & | \\ c_1 & c_2 & c_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} | \\ c_1 \\ | \end{bmatrix} - 1 \begin{bmatrix} | \\ c_2 \\ | \end{bmatrix} + 1 \begin{bmatrix} | \\ c_3 \\ | \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 15 \end{bmatrix}.$$

Conclusion

Both interpretations yield the same result:

$$B\mathbf{v} = \begin{bmatrix} 3 \\ 9 \\ 15 \end{bmatrix}.$$

The Transpose

Suppose that A is an m by n matrix. This means that A has m rows and n columns and is a map from $\mathbb{R}^n$ to $\mathbb{R}^m$.

The *transpose* map A^T is defined by: $A^T_{ij} = A_{ji}$. The transpose A^T is a map from $\mathbb{R}^m$ to $\mathbb{R}^n$. It has n rows and m columns.

$$
A = \begin{array}{cc} n & \\ \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} & m \end{array}
$$

$$
A^T = \begin{array}{cc} m & \\ \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & n \end{array}
$$

The *image* of A is the subspace of $\mathbb{R}^m$ consisting of vectors of the form Av where $v \in \mathbb{R}^n$. This is also called the *column space* of A and denoted by $\mathrm{Col}(A)$.

The *kernel* of A is the subspace of $\mathbb{R}^n$ consisting of vectors v such that $Av = 0$. This is also called the *null space* of A and denoted by $\mathrm{Nul}(A)$.

To re-iterate:

$$
\mathrm{Col}(A) \subseteq \mathbb{R}^m \qquad \mathrm{Nul}(A) \subseteq \mathbb{R}^n
$$

It is important to note that these two subspaces to *n*ot lie in the same ambiant space.

By similar reasoning the *image* of A^T is the subspace of $\mathbb{R}^n$ consisting of vectors of the form $A^T v$ where $v \in \mathbb{R}^m$. This is also called the *column space* of A^T and denoted by $\mathrm{Col}(A^T)$. This subspace is sometimes also called the *row space* of A.

The *kernel* of A^T is the subspace of $\mathbb{R}^m$ consisting of vecors v such that $A^T v = 0$. This is also called the *null space* of A^T and denoted by $\mathrm{Nul}(A^T)$. This subspace is sometimes called the *annihilator* of $\mathrm{Col}(A)$

We have:

$$\mathrm{Col}(A^T) \subseteq \mathbb{R}^n \qquad \mathrm{Nul}(A^T) \subseteq \mathbb{R}^m$$

Both $\mathrm{Col}(A^T)$ and $\mathrm{Nul}(A)$ lie in $\mathbb{R}^n$. And both $\mathrm{Col}(A)$ and $\mathrm{Nul}(A^T)$ lie in R^m

The main result of this section is the following:

Theorem 1. *The four fundamental subspaces satisfy the following orthogonal decompositions:*

$$\mathbb{R}^n = \mathrm{Col}(A^T) \oplus \mathrm{Nul}(A), \qquad \mathbb{R}^m = \mathrm{Col}(A) \oplus \mathrm{Nul}(A^\top)$$

where $\oplus$ denotes an orthogonal direct sum.

Proof. We must show that:

$$\boxed{\mathrm{Nul}(A) = \mathrm{Col}(A^T)^\perp} \quad \text{and} \quad \boxed{\mathrm{Nul}(A^\top) = \mathrm{Col}(A)^\perp}$$

Note that by symmetry we only really need to prove one of these identities. We will prove the one on the left.

Note firstly that the columns of A^T are simply the rows of A. So what we will actually prove is that:

$$\mathrm{Nul}(A) = \mathrm{Row}(A)^\perp$$

We shall show that a vector $\mathbf{x} \in \mathbb{R}^n$ satisfies $A\mathbf{x} = \mathbf{0}$ if and only if $\mathbf{x}$ is orthogonal to every vector in the row space of A.

Let the rows of A be $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m \in \mathbb{R}^n$. Then:

$$A\mathbf{x} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \mathbf{r}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{x} \end{bmatrix}$$

Thus,

$$A\mathbf{x} = \mathbf{0} \iff \mathbf{r}_i \cdot \mathbf{x} = 0 \quad \text{for all } i = 1, \ldots, m$$

Any vector $\mathbf{v} \in \text{Row}(A)$ is a linear combination:

$$\mathbf{v} = c_1\mathbf{r}_1 + \cdots + c_m\mathbf{r}_m$$

Then:

$$\mathbf{v} \cdot \mathbf{x} = \sum_{i=1}^{m} c_i(\mathbf{r}_i \cdot \mathbf{x}) = 0$$

So $\mathbf{x} \perp \mathbf{v}$, hence $\mathbf{x} \in \text{Row}(A)^\perp$.

Conversely, if $\mathbf{x} \in \text{Row}(A)^\perp$, then in particular $\mathbf{x} \perp \mathbf{r}_i$ for each row $\mathbf{r}_i$, so $A\mathbf{x} = \mathbf{0}$, meaning $\mathbf{x} \in \text{Nul}(A)$.

Therefore:

$$\boxed{\text{Nul}(A) = \text{Row}(A)^\perp}$$

$\square$

The Dual Space

Hopefully the reader has seen the idea of the *dual space* before. Let V be a finite dimensional vector space over the real numbers $\mathbb{R}$ The **dual space** of V, denoted V^*, is the set of all **linear functionals** on V. That is,

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

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 V^* is itself a vector space over $\mathbb{F}$, 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 Properties (Finite-Dimensional Case)

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

Proof. Let V be a finite-dimensional vector space over a field $\mathbb{R}$, and let $\dim V = n$. Choose a basis $\{e_1, e_2, \ldots, e_n\}$ for V.

Define linear functionals $e^1, e^2, \ldots, e^n \in V^* = \mathrm{Hom}(V, \mathbb{R})$ by

$$e^i(e_j) = \delta^i_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \quad \text{for all } 1 \leq i, j \leq n.$$

We claim that $\{e^1, \ldots, e^n\}$ is a basis for V^*.

Linear independence: Suppose $\sum_{i=1}^n a_i e^i = 0$ in V^*, where $a_i \in \mathbb{R}$. Applying both sides to e_j gives

$$0 = \left(\sum_{i=1}^n a_i e^i\right)(e_j) = \sum_{i=1}^n a_i e^i(e_j) = \sum_{i=1}^n a_i \delta^i_j = a_j.$$

Thus $a_j = 0$ for all j, so the set $\{e^1, \ldots, e^n\}$ is linearly independent.

Spanning: Let $f \in V^*$ be arbitrary. Define scalars $c_i = f(e_i) \in \mathbb{R}$ for $i = 1, \ldots, n$, and consider the functional

$$g = \sum_{i=1}^{n} c_i e^i \in V^*.$$

For any basis vector e_j, we have

$$g(e_j) = \sum_{i=1}^{n} c_i e^i(e_j) = \sum_{i=1}^{n} c_i \delta^i_j = c_j = f(e_j).$$

Since f and g agree on a basis of V, they agree on all of V; hence $f = g$. Therefore, every $f \in V^*$ is a linear combination of $\{e^1, \ldots, e^n\}$, so this set spans V^*.

Since $\{e^1, \ldots, e^n\}$ is a basis for V^*, we conclude that

$$\dim V^* = n = \dim V.$$

$\square$

As a consequence of the above, if $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis for V, 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

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

In $\mathbb{R}^n$ and $\mathbb{R}^m$, the dot product identifies vectors with linear functionals:

$$\mathbf{y} \in \mathbb{R}^m \quad \longleftrightarrow \quad f_{\mathbf{y}}(\mathbf{z}) = \mathbf{y}^\top \mathbf{z} \in (\mathbb{R}^m)^*.$$

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.

1. 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.

2. Matrix Representation of adjoint

Choose bases:

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

Suppose the matrix of T with respect to $\mathcal{B}, \mathcal{C}$ is $A = (a_{ij})$, so Then:

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

More generally:

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

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

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

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 $B = A^\top$.

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

The Annihilator

The **annihilator** provides the abstract, coordinate-free foundation for the orthogonal relationships among the four fundamental subspaces. It explains *why* these subspaces come in perpendicular pairs—and reveals that this structure is not just a coincidence of matrices, but a deep property of vector spaces and duality.

1. The Four Fundamental Subspaces

For a matrix $A \in \mathbb{R}^{m\times n}$, the four fundamental subspaces are:

Subspace	Location	Description
$\mathrm{Col}(A)$	$\mathbb{R}^m$	Column space (range)
$\mathrm{Nul}(A^\top)$	$\mathbb{R}^m$	Left null space
$\mathrm{Row}(A) = \mathrm{Col}(A^\top)$	$\mathbb{R}^n$	Row space
$\mathrm{Nul}(A)$	$\mathbb{R}^n$	Null space (kernel)

The key orthogonal decompositions are:

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

with

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

2. What Is 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).

- W° is a subspace of V^*,
- $\dim(W^\circ) = \dim(V) - \dim(W)$.

Dimension Theorem for Annihilators

Let V be a finite-dimensional vector space over a field $\mathbb{R}$, and let $W \subseteq V$ be a subspace. The **annihilator** of W is defined as

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

We prove that W° is a subspace of V^* and that

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

Step 1: W° is a subspace of V^*. Clearly $0 \in W^\circ$. If $f, g \in W^\circ$ and $\alpha, \beta \in \mathbb{R}$, then for any $\mathbf{w} \in W$,

$$(\alpha f + \beta g)(\mathbf{w}) = \alpha f(\mathbf{w}) + \beta g(\mathbf{w}) = 0,$$

so $\alpha f + \beta g \in W^\circ$. Hence $W^\circ \leq V^*$.

Step 2: Choose a basis adapted to W. Let $\dim V = n$ and $\dim W = k$. Choose a basis

$$\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$$

for W, and extend it to a basis of V:

$$\mathcal{B} = \{\mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}.$$

Let $\mathcal{B}^* = \{\mathbf{w}^1, \dots, \mathbf{w}^k, \mathbf{v}^{k+1}, \dots, \mathbf{v}^n\}$ be the dual basis of V^*, so that

$$\mathbf{w}^i(\mathbf{w}_j) = \delta^i_j, \quad \mathbf{v}^i(\mathbf{v}_j) = \delta^i_j, \quad \mathbf{w}^i(\mathbf{v}_j) = 0, \quad \mathbf{v}^i(\mathbf{w}_j) = 0$$

for all valid indices.

Step 3: Characterize W° using the dual basis. Let $f \in V^*$. Write f in the dual basis:

$$f = \sum_{i=1}^{k} a_i \mathbf{w}^i + \sum_{j=k+1}^{n} b_j \mathbf{v}^j.$$

For any $\mathbf{w} \in W$, we have $\mathbf{w} = \sum_{i=1}^{k} c_i \mathbf{w}_i$, so

$$f(\mathbf{w}) = \sum_{i=1}^{k} a_i c_i.$$

Thus $f(\mathbf{w}) = 0$ for all $\mathbf{w} \in W$ if and only if $a_1 = \cdots = a_k = 0$.

Therefore,

$$W^\circ = \operatorname{span}\{\mathbf{v}^{k+1}, \dots, \mathbf{v}^n\}.$$

Step 4: Compute the dimension. The set $\{\mathbf{v}^{k+1}, \dots, \mathbf{v}^n\}$ is linearly independent and spans W°, so it is a basis. Hence

$$\dim(W^\circ) = n - k = \dim(V) - \dim(W).$$

3. Connecting Annihilators to the Four Subspaces

In $\mathbb{R}^n$ and $\mathbb{R}^m$, the dot product identifies vectors with linear functionals:

$$\mathbf{y} \in \mathbb{R}^m \quad \longleftrightarrow \quad f_{\mathbf{y}}(\mathbf{z}) = \mathbf{y}^\top \mathbf{z} \in (\mathbb{R}^m)^*.$$

Under this identification:

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

Annihilators and Null Spaces via the Dot Product

In $\mathbb{R}^n$ and $\mathbb{R}^m$, the standard dot product identifies vectors with linear functionals:

$$\mathbf{y} \in \mathbb{R}^m \quad \longleftrightarrow \quad f_{\mathbf{y}} \in (\mathbb{R}^m)^*, \qquad f_{\mathbf{y}}(\mathbf{z}) = \mathbf{y}^\top \mathbf{z} = \mathbf{y} \cdot \mathbf{z}.$$

This is an isomorphism $\mathbb{R}^m \xrightarrow{\sim} (\mathbb{R}^m)^*$, since the dot product is nondegenerate.

Let A be an $m \times n$ real matrix. We prove:

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

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

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

Recall:

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

Its annihilator (in $(\mathbb{R}^m)^*$) is

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

Under the dot product identification, a functional $f \in (\mathbb{R}^m)^*$ corresponds to a unique vector $\mathbf{v} \in \mathbb{R}^m$ such that $f(\mathbf{y}) = \mathbf{v}^\top \mathbf{y}$ for all $\mathbf{y} \in \mathbb{R}^m$.

Thus, $\mathbf{v} \in \mathbb{R}^m$ corresponds to an element of $(\mathrm{Col}(A))^\circ$ iff

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

But $\mathbf{v}^\top A\mathbf{x} = (A^\top \mathbf{v})^\top \mathbf{x}$, so this holds for all $\mathbf{x}$ iff $A^\top \mathbf{v} = \mathbf{0}$.

Hence,

$$\mathbf{v} \in \mathrm{Nul}(A^\top) \quad \Longleftrightarrow \quad f_{\mathbf{v}} \in (\mathrm{Col}(A))^\circ.$$

Therefore, the dot product isomorphism restricts to an isomorphism

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

2. Proof that $\mathrm{Nul}(A) \cong (\mathrm{Row}(A))^\circ$.

Note that $\mathrm{Row}(A) = \mathrm{Col}(A^\top) \subseteq \mathbb{R}^n$. Applying the previous result to $A^\top$ (which is $n \times m$), we get:

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

where the annihilator is now in $(\mathbb{R}^n)^*$.

Alternatively, argue directly: $\mathbf{x} \in \mathbb{R}^n$ corresponds to $f_{\mathbf{x}} \in (\mathbb{R}^n)^*$ via $f_{\mathbf{x}}(\mathbf{z}) = \mathbf{x}^\top \mathbf{z}$. Then $f_{\mathbf{x}} \in (\mathrm{Row}(A))^\circ$ iff $f_{\mathbf{x}}(\mathbf{r}) = 0$ for every row vector $\mathbf{r}$ of A (viewed as elements of $\mathbb{R}^n$). Since any $\mathbf{r} \in \mathrm{Row}(A)$ is a linear combination of the rows of A, this is equivalent to

$$\mathbf{r}_i \mathbf{x} = 0 \quad \text{for each row } \mathbf{r}_i \text{ of } A,$$

which is precisely $A\mathbf{x} = \mathbf{0}$, i.e., $\mathbf{x} \in \mathrm{Nul}(A)$.

Thus, $\mathrm{Nul}(A) \cong (\mathrm{Row}(A))^\circ$.

Conclusion. Via the standard dot product identification $\mathbb{R}^k \cong (\mathbb{R}^k)^*$, we have natural isomorphisms:

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

4. The Big Picture: Duality

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:

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

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

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

5. Significance

1. **Unification**: The annihilator explains both orthogonal decompositions as instances of a single principle: the kernel of a linear map and the annihilator of its image are naturally paired.
2. **Coordinate-free insight**: The orthogonal relationships hold in any finite-dimensional inner product space—not just for matrices.
3. **Generalization**: In infinite-dimensional spaces (e.g., Hilbert spaces), the same duality holds via the Riesz representation theorem.
4. **Conceptual clarity**: The "four subspaces" are really two pairs of dual objects:
5. Domain side: $\mathrm{Row}(A)$ and $\mathrm{Nul}(A)$,
6. Codomain side: $\mathrm{Col}(A)$ and $\mathrm{Nul}(A^\top)$.

Conclusion

The annihilator is the abstract mechanism that explains why the four fundamental subspaces form orthogonal complements. It reveals that:

$$\mathrm{Nul}(A) \text{ is the annihilator of } \mathrm{Row}(A), \quad \mathrm{Nul}(A^\top) \text{ is the annihilator of } \mathrm{Col}(A).$$

Without the concept of the annihilator, these orthogonal relationships appear as computational coincidences; with it, they emerge as inevitable consequences of linear duality.

Theorem (Rank–Nullity Theorem)

Let $T : V \to W$ be a linear transformation between finite-dimensional vector spaces over a field $\mathbb{F}$. Then

$$\dim(V) = \dim(\ker T) + \dim(\operatorname{Im} T).$$

In particular, for any $m \times n$ matrix A over $\mathbb{F}$, viewing A as a linear map $\mathbb{F}^n \to \mathbb{F}^m$, we have

$$n = \text{nullity}(A) + \text{rank}(A).$$

Proof

Let $K = \ker T \subseteq V$. Since V is finite-dimensional, so is K. Let

$$\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$$

be a basis for K, so $k = \dim(\ker T)$.

Extend this to a basis for all of V:

$$\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_r\},$$

so that $\dim(V) = k + r$.

We claim that the set

$$\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_r)\}$$

is a basis for $\operatorname{Im} T$.

1. Spanning

Let $\mathbf{w} \in \operatorname{Im} T$. Then $\mathbf{w} = T(\mathbf{x})$ for some $\mathbf{x} \in V$. Write

$$\mathbf{x} = \sum_{i=1}^{k} a_i \mathbf{u}_i + \sum_{j=1}^{r} b_j \mathbf{v}_j.$$

Applying T, and using $T(\mathbf{u}_i) = \mathbf{0}$ (since $\mathbf{u}_i \in \ker T$), we get

$$T(\mathbf{x}) = \sum_{j=1}^{r} b_j T(\mathbf{v}_j).$$

Thus, $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_r)\}$ spans $\operatorname{Im} T$.

2. Linear Independence

Suppose

$$\sum_{j=1}^{r} c_j T(\mathbf{v}_j) = \mathbf{0}.$$

Then

$$T\left(\sum_{j=1}^{r} c_j \mathbf{v}_j\right) = \mathbf{0},$$

so $\sum_{j=1}^{r} c_j \mathbf{v}_j \in \ker T = \operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$.

But the full set $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_r\}$ is linearly independent. Therefore, the only linear combination of the $\mathbf{v}_j$'s that lies in $\operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is the trivial one. Hence, $c_j = 0$ for all j, and the set $\{T(\mathbf{v}_1), \dots, T(\mathbf{v}_r)\}$ is linearly independent.

3. Conclusion

We have shown that $\dim(\operatorname{Im} T) = r$. Since $\dim(V) = k + r$, it follows that

$$\dim(V) = \dim(\ker T) + \dim(\operatorname{Im} T).$$

This completes the proof.

Remarks

- This proof uses no matrices, coordinates, or row reduction.
- It works over any field (e.g., $\mathbb{R}$, $\mathbb{C}$, finite fields).
- It reveals that Rank–Nullity is a structural property of linear maps, not a computational artifact.

Introduction

A fundamental fact in linear algebra is that the rank of a matrix A is equal to the rank of its transpose $A^\top$. In other words, the maximum number of linearly independent **columns** of A is the same as the maximum number of linearly independent **rows** of A.

This might seem surprising at first—after all, rows and columns live in different spaces! But in $\mathbb{R}^n$ (or $\mathbb{C}^n$), the presence of an inner product (the dot product) creates a deep symmetry between rows and columns.

We will prove this result using two key ideas:

1. The **Rank–Nullity Theorem**,
2. The **orthogonal relationship** between the null space and the row space.

This proof works for real matrices (and complex matrices with minor adjustments), but it relies on the geometry of the dot product.

Step 1: What the Rank–Nullity Theorem Tells Us

Let A be an $m \times n$ real matrix. We can think of A as a linear transformation that maps vectors from $\mathbb{R}^n$ to $\mathbb{R}^m$.

The Rank–Nullity Theorem says:

(dimension of domain) = (dimension of null space)+(dimension of column space).

In symbols:

$$n = \dim(\mathrm{Nul}(A)) + \dim(\mathrm{Col}(A)). \tag{1}$$

Here:

- $\mathrm{Nul}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$ is the set of vectors that A sends to zero,
- $\mathrm{Col}(A)$ is the span of the columns of A, and its dimension is the **column rank** of A.

Step 2: The Geometric Link Between Rows and Null Space

Now consider the **row space** of A, denoted $\mathrm{Row}(A)$. This is the subspace of $\mathbb{R}^n$ spanned by the rows of A.

Here's the key geometric insight:

A vector $\mathbf{x}$ is in the null space of A if and only if it is perpendicular to every row of A.

Why? Because the equation $A\mathbf{x} = \mathbf{0}$ means that the dot product of each row of A with $\mathbf{x}$ is zero.

In other words:

$$\mathrm{Nul}(A) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \perp \mathbf{r} \text{ for every row } \mathbf{r} \text{ of } A\}.$$

This is precisely the definition of the **orthogonal complement** of the row space. So we have:

$$\mathrm{Nul}(A) = \mathrm{Row}(A)^{\perp}. \tag{2}$$

Step 3: Use Dimensions of Orthogonal Complements

In $\mathbb{R}^n$, if S is any subspace, then:

$$\dim(S) + \dim(S^{\perp}) = n.$$

Apply this to $S = \mathrm{Row}(A)$. Using (2), we get:

$$\dim(\mathrm{Row}(A)) + \dim(\mathrm{Nul}(A)) = n. \tag{3}$$

But $\dim(\mathrm{Row}(A))$ is exactly the **row rank** of A.

Step 4: Compare with Rank–Nullity

Now look back at equation (1) from Rank–Nullity:

$$n = \dim(\mathrm{Nul}(A)) + \dim(\mathrm{Col}(A)).$$

And equation (3) from orthogonality:

$$n = \dim(\mathrm{Nul}(A)) + \dim(\mathrm{Row}(A)).$$

Both right-hand sides equal n, and both contain the term $\dim(\text{Nul}(A))$. Therefore, the remaining terms must be equal:

$$\dim(\text{Col}(A)) = \dim(\text{Row}(A)).$$

In other words:

$$\text{column rank of } A = \text{row rank of } A.$$

Since the row rank of A is the same as the column rank of $A^\top$, we conclude:

$$\text{rank}(A) = \text{rank}(A^\top).$$

Conclusion

The equality of row and column rank is not a coincidence—it is a consequence of the **geometric structure** of Euclidean space. The dot product creates a perfect pairing between the row space and the null space, and the Rank–Nullity Theorem translates this geometric fact into an algebraic equality of dimensions.

This proof beautifully ties together:

- Linear algebra (Rank–Nullity),
- Geometry (orthogonality),
- Matrix theory (rows vs. columns).