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.

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

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

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 $\mathrm{Nul}(A^T)$.

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 $\mathbf{v} \in \mathrm{Row}(A)$ 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:

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

$\square$

Theorem (Rank–Nullity Theorem)

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

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

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

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

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

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 $\dim(\mathrm{Nul}(A))$. Therefore, the remaining terms must be equal:

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

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:

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