THE FUNDAMENTAL THEOREM OF LINEAR ALGEBRA

Theorem 0.1. *For any $m \times n$ matrix A the dimension of the Column Space and the dimension of the Null Space add up to the total number of columns, i.e.,*

$$\dim C(A) + \dim N(A) = n \ .$$

Proof. We denote $r = \dim C(A)$, the rank of the matrix and $k = \dim N(A)$. The idea is to construct a *basis* of the form $v_1, \dots, v_r, u_1, \dots u_k$ for $\mathbb{R}^n$. It then follows that $k + r = n$. Pick a basis $w_1, \dots, w_r$ in $C(A)$. Since these vectors are in the column space of A there exists vectors $v_1, \dots, v_r$ in $\mathbb{R}^n$ so that $w_j = Av_j, j = 1, \dots, r$. For the remaining vectors we pick a basis for the null space $u_1, \dots, u_k$. We have to show that $v_1, \dots, v_r, u_1, \dots, u_k$ is a basis for $\mathbb{R}^n$, which shows that $r + k = n$.

To establish this we have to prove that these vectors span $\mathbb{R}^n$ and that they are linearly independent. For the independence we write

$$0 = \sum_{j=1}^{r} \alpha_j v_j + \sum_{\ell=1}^{k} \beta_\ell u_\ell \tag{1}$$

and want to show that the coefficients $\alpha_1 = \cdots = \alpha_r = \beta_1 = \cdots = \beta_k = 0$. Apply the matrix A to (1) to get

$$0 = \sum_{j=1}^{r} \alpha_j w_j$$

where we have used that the u_js are in $N(A)$ and that $Av_j = w_j, j = 1, \dots r$. Since the w_js are independent the α_js are all zero. But then, by (1), since the u_js are also independent we must have that the β_js vanish also. Next we show that $w_1, \dots, w_r, u_1, \dots, u_k$ span $\mathbb{R}^n$. Pick any vector $x \in \mathbb{R}^n$ and apply A to it. We have that $Ax \in C(A)$, and since $w_1, \dots, w_r$ is a basis for $C(A)$

$$Ax = \sum_{j=1}^{r} c_j w_j \ .$$

This shows that $x - \sum_{j=1}^{r} c_j v_j \in N(A)$ which is spanned by the u_js and this establishes the claim. □

What is interesting is that we did not use any row reduction to prove this theorem. All we needed was that the spaces involved are finite dimensional and the existence of a basis.

The next theorem connects the matrix A with its transpose A^T.

Theorem 0.2. *We have*

$$\dim C(A^T) = \dim C(A) = \text{rank}(A) \ .$$

In other words the row vectors and column vectors span spaces of the same dimension.

Proof. We argue with row reduction. Bringing A into row reduced echelon form, the number of rows with a pivot and the number of columns with a pivot must be the same. This number is the rank of the matrix. The rows with a pivot are linearly independent and span the row space of the row reduced matrix. Since the row space did not change under row operations this is also a basis for the row space of the matrix A. Hence $\dim A = \dim A^T = \operatorname{rank} A$. ☐

There is a strong connection between the the transpose matrix and the dot product. The following is completely elementary.

Lemma 0.3. *Let A be an $m \times n$ matrix and $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. Then*

$$y^T A x = (A^T y)^T x \ . \tag{2}$$

If x is perpendicular to all vectors in $C(A^T)$ then $(A^T y)^T x = 0$ for all $y \in \mathbb{R}^m$ and by the above formula $y^T A x = 0$ for all $y \in \mathbb{R}^m$. Hence $Ax = 0$ and $x \in N(A)$. To understand this in a better way, we introduce the orthogonal complement.

Definition 0.4. *Let V be a subspace of $\mathbb{R}^n$. The orthogonal complement of V in $\mathbb{R}^n$ is given by*

$$V^\perp := \{x \in \mathbb{R}^n : x^T y = 0, \text{ all } y \in V\}$$

It is an exercise to convince yourself that $V^\perp$ is also a subspace of $\mathbb{R}^n$. Using this notion of orthogonal complement we obtain the important relation

$$C(A^T)^\perp = N(A) \ . \tag{3}$$

The key result about this is

Theorem 0.5. *Every vector $x \in \mathbb{R}^n$ can be written in a unique way as a sum of two vectors, $y \in V$ and $z \in V^\perp$.*

Proof. Here is the idea: Pick any $x \in \mathbb{R}^n$ and try to find a vector $y \in V$ that is closest to x. Suppose for the moment that such a vector y exists. We claim that $x - y$ must be perpendicular to all the vectors in V. To see this pick any $v \in V, t \in \mathbb{R}$ and consider the function

$$f(t) := \|x - (y + tv)\|^2 \ .$$

Note that $y + tv \in V$ since V is a subspace. Hence $f(t) \geq f(0)$ because we assumed that y is the vector in V that is closest to x. Now we compute

$$f(t) = \|x - y\|^2 + 2t(x - y)^T v + t^2 \|v\|^2 \ .$$

Obviously $f(t)$ is differentiable and since it has a minimum at $t = 0$ we must have that $f'(0) = 0$. But this means that $(x - y)^T v = 0$. Since $v \in V$ is arbitrary we learn that $x - y \perp V$. In other words if we set $x - y = z$

$$x = z + y \ , \ y \in V \ , \ z \in V^\perp \ .$$

This decomposition is unique because if $x = y_1 + z_1$ then

$$y + z = y_1 + z_1 \text{ or } y - y_1 = z_1 - z$$

which means that $y - y_1 \in V \cap V^\perp$ and thus $y = y_1$.

The problem that remains to be solved is to find y or at least show that it exists. ☐

In particular

$$[V^\perp]^\perp = V \ .$$

This means, that the only vector in $\mathbb{R}^n$ that is perpendicular to V as well as $V^\perp$ is the zero vector.