By looking at the reduced row echelon form of a matrix, we have determined that, for any matrix A,

(1) $\dim(\text{row}A) \;=\; \dim(\text{col}A)$

(2) $\dim(\text{col}A) \;+\; \dim(\text{row}A) \;=\;$ number of columns of A

The dimension of the row space of A is sometimes called the *row rank* of A, and the dimension of the column space of A is sometimes called the *column rank* of A. Since they are, in fact, the same, it is not ambiguous to refer simply to the *rank* of a matrix. The dimension of the null space of A is called the *nullity* of A. The first statement is sometimes called the "Rank Theorem," and the second the "Rank-Nullity Theorem."

We will discuss these results now from the perspective of linear transformations.

1. Kernel and Image

Recall that a *linear transformation* is a function $T : \mathbb{R}^n \to R^m$ such that

(1) $T(\alpha\, v) \;=\; \alpha\, T(v)$ for all scalars $\alpha \in \mathbb{R}$ and all vectors $v \in \mathbb{R}^n$
(2) $T(v + w) \;=\; T(v) \;+\; T(w)$ for all vectors $v, w \in \mathbb{R}^n$

Recall that the set of *all* vectors whose image is the zero vector is called the *kernel* of T:

$$\ker T \;=\; \{\text{vectors } v \text{ such that } T(v) = 0\}$$

The fact that $T(0)$ is always zero implies that the zero vector is always in the kernel. When the kernel consists *only* of the zero vector, we say that the kernel is *trivial*.

Recall that the *image* of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is the collection of all vectors $w \in \mathbb{R}^m$ such that $Tv = w$ for some $v \in \mathbb{R}^n$.

Recall that the matrix for a linear transformation T is matrix whose columns are obtained by evaluating T on the standard basis elements $e_1, \ldots e_n$ of $\mathbb{R}^n$:

$$\left(\begin{array}{c|c|c|c} Te_1 & Te_2 & \ldots & Te_n \end{array}\right)$$

Thus the kernel of a linear transformation is the *null space* of its associated matrix, and the image of a linear transformation is the *column space* of its associated matrix. What is the row space? Well, since the rows of a matrix are the columns of its transpose, the row space of a matrix is the image of the linear transformation associated to the transpose matrix. In summary, if A is a (matrix for a) linear transformation,

$$\ker A \;=\; \text{null}(A) \qquad \text{Img}A \;=\; \text{col}(A) \qquad \text{Img}A^T \;=\; \text{row}(A)$$

2. Complementarity of kernel and image

Statement (2) above, i.e. that the dimension of the column space of a matrix plus the dimension of the null space of the matrix is equal to the number of columns of the matrix, follows from a certain complementarity of the kernel and the image of a linear transformation, which we discussed in the case of linear transformations of the plane.

Recall that rotations, reflections, dialations, and shears of the plane all had trivial kernel and image equal to the whole plane. In the case of a projection onto a line through the origin, the image was that line and the kernel was the perpendicular line through the origin. Finally, the zero map has image zero and kernel equal to the whole plane. In all of these cases, notice that the kernel and the image are orthogonal to each other and that they span the whole plane.

In general, for a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, the kernel of T is a subspace of the source $\mathbb{R}^n$, and the image of T is a subspace of the target $\mathbb{R}^m$, and further, the kernel and the image are complementary, in a sense that can be made precise:

$$\text{“}\ \mathbb{R}^n \ \approx \mathrm{ker}T \ \oplus \ \mathrm{Img}T\ \text{”}$$

We will not discuss the precise meaning of this right now. Roughly, this means that the kernel and the image together make up a copy of $\mathbb{R}^n$. In any case, the linear transformation version of statement (2) follows from this complementarity:

$$n \ = \ \dim\big(\mathrm{ker}T\big) \ + \ \dim\big(\mathrm{Img}T\big)$$

Notice that this means that the image of a linear transformation has maximal dimension if and only if the kernel is trivial.

Exercise 1. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Prove that the kernel of T is a subspace of $\mathbb{R}^n$, and the image of T is a subspace of $\mathbb{R}^m$.

3. Classification of solutions of linear systems

A consequence of this complementarity is the classification of solutions of linear systems.

Theorem (Classification of solutions of linear systems)**.** *A linear system either (1) has no solutions, (2) has a unique solution, or (3) has infinitely many solutions*

Proof. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $w \in \mathbb{R}^m$. It suffices to show that if there are two or more distinct solutions to the equation $Tv = w$, then there are in fact infinitely many solutions. We will use the following Lemma:

Lemma. *Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. If $Tv = w$ for some $v \in \mathbb{R}^n$ and some $w \in \mathbb{R}^m$, then $T(v + u) = w$ for all $u \in \mathrm{ker}T$.*

whose proof is relegated to an exercise.
Suppose v_1 and v_2 are distinct solutions, i.e. $Tv_1 = Tv_2 = w$, but $v_1 \neq v_2$. Consider $u = v_1 - v_2$. Then

$$Tu \ = \ T(v_1 - v_2) \ = \ Tv_1 - Tv_2 \ = \ w - w \ = \ 0$$

i.e. $u = v_1 - v_2$ is in the kernel. Since $v_1 \neq v_2$, this means that the kernel is nontrivial, i.e. it contains more than just the zero vector. Since the kernel is a subspace, it also contains all scalar multiples of u. (Infinitely many, since $u \neq 0$!)

By the lemma, $T(v_1 + \alpha u) = w$ for all scalars α. We have demonstrated that there are infinitely many solutions to $Tv = w$, i.e. all vectors of the form $v_1 + \alpha u$, where α is a scalar. □

Exercise 2. Prove the lemma used in the proof of the theorem above.

4. Duality: transposes and the dot product

To discuss statement (1) from the perspective of linear transformations, we need a deeper understanding of the transpose of a matrix.

Our initial definition of the transpose of a matrix is that it is the matrix obtained by switching rows and columns. What does this mean in terms of linear transformations? A $(m \times n)$ matrix A is a linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$, so its transpose, being of size $(n \times m)$ is a transformation from $\mathbb{R}^m \to \mathbb{R}^n$. The image of A^T is a subspace of $\mathbb{R}^n$.

For a linear transformation with matrix A, the transpose matrix is characterized by:

$$Au \cdot w \;=\; u \cdot A^T w$$

Remember that the dot product encapsulates the *geometry* of vectors, since it defines lengths and angles. Thus we can interpret this relation as a description of how the linear transformation interacts with the geometry.

In fact, any vector in $\mathbb{R}^n$ can be written as a sum of a vector in $\ker A$ and a vector in $\mathrm{Img} A^T$, and any vector in the kernel of A is orthogonal to any vector in the image of A^T. This is summarized with the following notation:

$$\mathbb{R}^n \;=\; \ker A \,\oplus\, \mathrm{Img} A^T$$

Exercise 3. Consider the example matrix A from last week's in-class activity, and show that any vector in $\ker A$ ($= \mathrm{null}(A)$) is orthogonal to any vector in $\mathrm{Img} A^T$ ($\mathrm{row}(A)$).

In any case, this implies that

$$n \;=\; \dim\big(\ker A\big) \,+\, \dim\big(\mathrm{Img} A^T\big)$$

i.e.

$$n \;=\; \dim\big(\mathrm{null}(A)\big) \,+\, \dim\big(\mathrm{row}(A)\big)$$

Combining this with the (2), this implies that

$$\dim\big(\mathrm{Img} A\big) \;=\; \dim\big(\mathrm{Img} A^T\big)$$

i.e.

$$\dim\big(\mathrm{col}(A)\big) \;=\; \dim\big(\mathrm{row}(A)\big)$$

This is sometimes called the "Rank Theorem."

5. Invertible Matrices

For a linear transformation to have a well defined inverse, it certainly must be one-to-one.

Lemma. *A linear transformation is one-to-one if and only if its kernel is trivial.*

Exercise 4. Prove the lemma above.

Notice that the complementarity result

$$n \;=\; \dim\big(\ker T\big) \,+\, \dim\big(\mathrm{Img} T\big)$$

implies that the image of a linear transformation has maximal dimension if and only if the kernel is trivial.

It turns out that, for linear transformations $\mathbb{R}^n \to \mathbb{R}^n$ (which have square matrices) to be invertible, it suffices for the kernel to be trivial. By the remark above, this is equivalent to the image being the whole of $\mathbb{R}^n$, i.e. the linear transformation is *onto*.

6. Linear Systems and Linear Dependence of Vectors

We now discuss the implications of these results for linear systems and for tests for linear dependence.

Let A be an $(n \times n)$ matrix. If A is invertible, then ...

Consider the linear system corresponding to the matrix equation $Ax = b$. Since A is onto, there will always be a solution (i.e. the system is consistent). Since A is one-to-one, the solution will always be unique (i.e. the linear equations are independent.) And, in fact, the solution is $x = A^{-1}b$.

From the column vector perspective, this is telling us that any $b \in \mathbb{R}^n$ can be written uniquely as a linear combination of the columns of A.

Consider the homogeneous linear system corresponding to the matrix equation $Ax = 0$. The zero vector is always a solution. Since A is one-to-one, it is the unique solution.

This means that the kernel/null space of A is trivial, so the nullity of A is zero.

From the column vector perspective, this is telling us that the only way to write 0 as a linear combination of the columns of A is to have all the coefficients be zero, i.e. the columns of A are linearly independent. This means that the columns of A form a basis for the column space of A, and that the (column) rank of A is n.

By duality, A^T is also invertible, so its columns are linearly independent. This implies that the rows of A are linearly independent, so they form a basis for the row space of A, and the (row) rank of A is n.

If, on the other hand, A is not invertible, that means that the kernel/null space is nontrivial (so the nullity is nonzero) and the image is not maximal (so the rank is less than n). Thus an equation $Ax = b$ may or may not have a solution (since A is not onto), and if it does have a solution, the solution will not be unique (because if v is a solution $v + u$ is a solution for all u in the kernel). A vector b may or may not be able to be written as a linear combination of the columns of A. The homogeneous equation $Ax = 0$ has nontrivial solutions. The columns of A are linearly dependent. By duality, A^T is not invertible, so the columns of A^T and thus the rows of A are linearly dependent.