Table of Contents and Results

1. Common Matrices

Below, we give examples of some common matrices (in the 3x3 case) so as to have a reference.

Common in elimination:

Identity, $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Diagnol, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

Elimination, $E_{(2,1)} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Upper Triangular, $U = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & -7 \\ 0 & 0 & 1 \end{bmatrix}$.

Lower Triangular, $L = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix}$. Permutation, $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

Common linear transformations:

Shear, $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Rotation, $\begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

Scale, $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Reflection, $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

Projection, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Common in statistics or other disciplines:

Stochastic Matrices, $\begin{bmatrix} 0.05 & 0.2 & 0.5 \\ 0.4 & 0.2 & 0.5 \\ 0.55 & 0.6 & 0 \end{bmatrix}$. Each entry in (i, j) represents the probability of moving from state i to state j. The sum of each column is 1, since the total probability of moving to any state from a state must be 1. The long-term or stationary state of the matrix is found by taking higher and higher powers of the matrix.

Incidence matrices, $\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ -1 & 0 & 1 \end{bmatrix}$. The matrix can represent a directed graph. Each column represents the nodes of the graph and each row represents the edges of the graph. The positive or negative sign reflects the direction of each edge; if it is leaving a node it is negative and if it is heading toward a node it is positive.

2. Elimination

We have already explored how matrices can represent linear transformations between vector spaces. Now, we would like to look at matrices throught the lense of solving a system of linear equations. We have m equations and n unknowns. In general, when the number of equations is greater than the number of unknowns ($m > n$), one would expect no solutions to exist. On the other hand, when the number of unknowns is greater than the number of equations ($n > m$), one would expect many solutions to exist. We will spend most of our time focusing on the instances where the number of equations is equal to the number of unknowns. Such cases either have a unique solution, no solution, or infinitely many solutions.

There is an algorithmic way to solve a system of linear equations, called **Gaussian Elimlination**. For illustrative purposes, we take a generalized three-by-three matrix $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ where each entry in the matrix represents the coefficients in the linear combination. Our goal is to solve for the unknown $x = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix}$. Our constraint is $b = \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix}$, so the system is $Ax = b$. If we form an **augmented matrix** $A' = \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_{11} \\ a_{21} & a_{22} & a_{23} & b_{21} \\ a_{31} & a_{32} & a_{33} & b_{31} \end{bmatrix}$, then permuting the rows or adding multiples of one row to another do not change the system--exchanging rows one and two is equivalent to solving $\begin{bmatrix} x_{21} \\ x_{11} \\ x_{31} \end{bmatrix}$ with $\begin{bmatrix} a_{21} & a_{22} & a_{23} & b_{21} \\ a_{11} & a_{12} & a_{13} & b_{11} \\ a_{31} & a_{32} & a_{33} & b_{31} \end{bmatrix}$ while subtracting row one from row two is equivalent to solving $\begin{bmatrix} x_{21} \\ x_{11} - x_{21} \\ x_{31} \end{bmatrix}$ with $\begin{bmatrix} a_{11} & a_{12} & a_{13} & b_{11} \\ a_{21} - a_{11} & a_{22} - a_{12} & a_{23} - a_{13} & b_{21} - b_{11} \\ a_{31} & a_{32} & a_{33} & b_{31} \end{bmatrix}$. The idea of elimination is to utilize these two operations to simplify the system as much as possible. Notice that if all but one of the coefficients are eliminated from the system in a given row, we are left with a coeficient, a constraint, and an unknown, and can just divide the constraint by the coefficient to find the unknown. To achieve this form, we put the matrix in **upper-triangular form**, or **reduced echelon form** with 0's everywhere below the main diaganol. We call the enteries on the main diaganol the **pivots** of the matrix.

First we add $\frac{-a_{21}}{a_{11}}$ multiples of the first row from the second row to get

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & b_{11} \\ a_{21} & a_{22} & a_{23} & b_{21} \\ a_{31} & a_{32} & a_{33} & b_{31} \end{bmatrix} \xrightarrow[r_2*=r_2+\left(\frac{-a_{21}}{a_{11}}\right)r_1]{} \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_{11} \\ 0 & \left[a_{22}+\left(\frac{-a_{21}a_{12}}{a_{11}}\right)\right] & \left[a_{23}+\left(\frac{-a_{21}a_{13}}{a_{11}}\right)\right] & \left[b_{21}+\left(\frac{-a_{21}b_{11}}{a_{11}}\right)\right] \\ a_{31} & a_{32} & a_{33} & b_{31} \end{bmatrix}$$

. Next we make the second entry under the first pivot zero with $r_3 *= r_3 + \left(\frac{-a_{31}}{a_{11}}\right) r_1$. Finally, we make the first entry under the second pivot zero with $r_3 *= r_3 + \left(\frac{-a_{32}}{a_{12}}\right) r_1$. If at any time a pivot is zero, one can exchange the row with another if the matrix allows. We call this process **forward elimination** and the resultant matrix U for upper triangular. Back-substitution yields the solution.

An example may help clarify. Consider the following: three people go to the market to buy fruit. The first person buys an apple, a banana, and 3 oranges for \$8.50. The next person buys 3 apples, 2 bananas, and 2 oranges for \$12. The final person buys 4 bananas and an orange for \$9. The situation can be represented by a system of linear equations as shown below. The goal is to find, via matrix multiplication, the unit price of each fruit.

$$p_a + p_b + 3p_o = 8.5$$
$$3p_a + 2p_b + 2p_o = 10$$
$$4p_b + p_o = 8$$

So the known quantities can be assembled into a matrix, call it A. Likewise, the vector of unit prices can be assembled into a column matrix, call it $\vec{p}$. The product, $A\vec{p}$, will equal the matrix of total costs. We have the following:

$$\begin{bmatrix} 1 & 1 & 3 \\ 3 & 2 & 2 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} p_a \\ p_b \\ p_o \end{bmatrix} = \begin{bmatrix} 8.5 \\ 10 \\ 8 \end{bmatrix}$$

To solve for the solution set, we use Gaussian Elimination, and thus create the augmented matrix below.

$$\begin{bmatrix} 1 & 1 & 3 & 8.5 \\ 3 & 2 & 2 & 10 \\ 0 & 4 & 1 & 8 \end{bmatrix}$$

To follow along with the row operations, each step is shown along with the transformations of the matrix. Following convention, r_i represents the i^{th} row and r_i^*represents the new i^{th} column.

$$\begin{bmatrix} 1 & 1 & 3 & 8.5 \\ 3 & 2 & 2 & 10 \\ 0 & 4 & 1 & 8 \end{bmatrix} \xrightarrow[-3r_1 + r_2 = r_2^*]{} \begin{bmatrix} 1 & 1 & 3 & 8.5 \\ 0 & -1 & -7 & -15.5 \\ 0 & 4 & 1 & 8 \end{bmatrix} \xrightarrow[4r_2{}^* + r_3 = r_3^*]{} \begin{bmatrix} 1 & 1 & 3 & 8.5 \\ 0 & -1 & -7 & -15.5 \\ 0 & 0 & -27 & -54 \end{bmatrix}$$

With back-substitution, we see that $-27p_o = -54$, so $p_o = 2$.
We then see that $-1p_b - 7p_o = -15.5$, so $-1p_b - 7(2) = -15.5$, and thus $p_b = 1.5$.
Finally, we have $1p_a + 1p_b + 3p_o = 8.5$, so $1p_a + 1(1.5) + 3(2) = 8.5$, and thus $p_a = 1$.

It took two operations bring $A = \begin{bmatrix} 1 & 1 & 3 \\ 3 & 2 & 2 \\ 0 & 4 & 1 \end{bmatrix}$ to $U = \begin{bmatrix} 1 & 1 & 3 \\ 0 & -1 & -7 \\ 0 & 0 & -27 \end{bmatrix}$. The operations that brought the matrix to U can be represented with matrix multiplication. We call the matrices which perform these operation **Elementary Matrices**, and denote them $E_{(i,j)}$ where i represents which row is changing and j represents which row is being used in the elimination process. Sticking with the example, the first operation created a 0 in the 2nd row and 1st column of A. So we are looking for the $E_{(2,1)}$ matrix such that $E_{(2,1)}A = A_1^*$ where A_1^* represents the matrix A after the first row operation.

$$E_{(2,1)} \begin{bmatrix} 1 & 1 & 3 \\ 3 & 2 & 2 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & -1 & -7 \\ 0 & 4 & 1 \end{bmatrix}$$

We should pause to consider the various ways in which one can imagine matrix multiplication. For any matrices A and B which permit $AB = C$, we can compute the exact value of a single entry in the $n \times n$ matrix C as $C_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$, where a_{ik} is the element in the i^{th} row and k^{th} column of A while b_{ik} is the element in the i^{th} row and k^{th} column of B. This is useful to find individual entries, but less natural than solving for whole rows or columns at a time. The **column view** of multiplication shows the j^{th} column of C as a product of A and a matrix of the j^{th} column of B, $AB_j = C_j$. In the case where B is a column matrix, the resultant matrix is the linear combinations of the entries in B and the columns in A. Likewise the **row view** of multiplication shows the i^{th} row of C as a product of the i^{th} row of A and the full matrix B, $A_iB = C_i$.

Returning to the example, the matrix which allows the multiplication $E_{(2,1)}A = A_1^*$ is the identity matrix in all entries but the 2nd row and first column. This entry is the multiple of row 1 which is being added to row 2, in this case -3. So $E_{(2,1)} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The same process holds in general-- $E_{(i,j)}$ is the identity matrix with the multiple of row j that is added to row i in the entry A_{ij}. In our example, we had one more transformation, taking A_1^* to U. So the goal is to solve for $E_{(3,2)}$ so that $E_{(3,2)}A_1^* = U$. By the same process as described above, we get $E_{(3,2)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}$. See that since $E_{(2,1)}A = A_1^*$ and $E_{(3,2)}A_1^* = U$, $E_{(3,2)}\left(E_{(2,1)}A\right) = U$. Since matrix multiplication is associative (but not commutative!) we have that $\left(E_{(3,2)}E_{(2,1)}\right)A = U$. This is convenient because we can represent the entire transformation from A to U, the upper triangular, with one elimination matrix E. In the example, we see that $E = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 4 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\-3 & 1 & 0\\0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\-3 & 1 & 0\\-12 & 4 & 1\end{bmatrix}$, and can verify $EA = U$.

In what may first seem counter-intuitive, we would instead like to write A as a product of matrices, $A = E^{-1}U$. Inverting E amounts to doing the steps of elimination in reverse. If $E_{(3,2)}E_{(2,1)} = E$, then surely ${E_{(2,1)}}^{-1}{E_{(3,2)}}^{-1} = E^{-1}$ since $E^{-1}E = E^{-1}\left(E_{(3,2)}E_{(2,1)}\right) = \left({E_{(2,1)}}^{-1}{E_{(3,2)}}^{-1}\right)\left(E_{(3,2)}E_{(2,1)}\right) = I$. In our case, we see that ${E_{(2,1)}}^{-1} = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & 0 & 1\end{bmatrix}$ and ${E_{(3,2)}}^{-1} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & -4 & 1\end{bmatrix}$ so $E^{-1} = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & -4 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & -4 & 1\end{bmatrix}$. We call E^{-1} the matrix L for **lower-triangular.** This matrix is interesting for two reasons. Most obviously, its shape is opposite U with 0's everywhere above the main diagonal. Secondarily, note that it completely contains the inverse of the multiples that brought A to U! We see 3 in the second row and first column, and we subtracted 3 of the first row from the second to obtain a new second row. Likewise, we see -4 in the third row and second column, and we add four of the second row to the third to obtain a new third row. It is nice to see $A = \begin{bmatrix}1 & 1 & 3\\3 & 2 & 2\\0 & 4 & 1\end{bmatrix} = LU = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & -4 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 3\\0 & -1 & -7\\0 & 0 & -27\end{bmatrix}$.

There is one major difference between L and U besides their obvious differences in makeup. See that the main diagonal of L is filled with 1's while the main diagonal of U contains the pivots. If we write U as a product of matrices, then we can pull the pivots out of U and have 1's on the main diagonal of both L and U. We remark without proof that the product of the pivots is the determinant of the matrix! Call the matrix of pivots D for the **diagonal matrix**. Then we can write $A = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & -4 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & -1 & 0\\0 & 0 & -27\end{bmatrix}\begin{bmatrix}1 & 1 & 3\\0 & 1 & -7\\0 & 0 & 1\end{bmatrix} = LDU$. This is just about optimal. However there is one issue with this factorization. Note that if a zero appears at any time in the pivot position, we cannot always eliminate the other entries below the pivot. There is an easy solution, which is to exchange the rows of the matrix prior to factoring. We are looking for the matrix P (for permutation) which takes A to A^*. We have A^* in mind already, it is simply A with a reordering of rows. The entries along each i^{th} row in P are the multiples of each row in A which are added to achieve the i^{th} row of A^*. If we want the first row of A^* to be the second row of A, then the first row of P will be $< 0,1, ... >$; if we want the third row of A^* to be the fourth row of A, then the third row of P will be $< 0,0,0,1, ... >$, etc. For a concrete example, imagine if our example originally had the third equation in place of the first. Then the coefficient matrix would be $\begin{bmatrix} 0 & 4 & 1 \\ 1 & 1 & 3 \\ 3 & 2 & 2 \end{bmatrix}$. To get a non-zero first pivot, we could achieve our original matrix with $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, $PA = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 4 & 1 \\ 1 & 1 & 3 \\ 3 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 3 & 2 & 2 \\ 0 & 4 & 1 \end{bmatrix}$. So the complete factorization is really $PA = LDU$, or $A = P^{-1}LDU$.

The ideas in Gaussian Elimination can be used for more then just solving systems of linear equations. We have seen in previous papers that it is generally computationally difficult to solve for A^{-1} with something like Cramer's Rule. Instead we may observe that since $AA^{-1} = I$, if we perform operations on A to turn it to I, then performing the same operations to I should yield A^{-1}. By bringing A to U, I goes to L^{-1} (recall L is calculated as the inverse of the elimination steps, so the regular elimination steps must be L^{-1}), then by bringing U to I, L^{-1} goes to A^{-1}. This is called the **Gauss-Jordan Method**, Which is a more strict form of elimination where all the values above and below the pivots are brought to zero and all the pivots are divided out to be 1.

Note that the columns of I are combinations of A and the columns in A^{-1}. We can create an augmented matrix $[AI] = \begin{bmatrix} 1 & 1 & 3 & 1 & 0 & 0 \\ 3 & 2 & 2 & 0 & 1 & 0 \\ 0 & 4 & 1 & 0 & 0 & 1 \end{bmatrix}$. We've already shown A goes to U with the operations $-3r_1 + r_2 = r_2^*$ and $4r_2{}^* + r_3 = r_3^*$. We are left with $[UL^{-1}] = \begin{bmatrix} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & -1 & -7 & -3 & 1 & 0 \\ 0 & 0 & -27 & -12 & 4 & 1 \end{bmatrix}$. We now want to create zeros above the main diagonal. The operations $\frac{-7}{27}r_3 + r_2 = r_2{}^*$, $\frac{3}{27}r_3 + r_1 = r_1{}^*$, and $r_2 + r_1{}^* = r_1{}^{**}$ get us to the diagonal of pivots, then dividing the second row by -1 and the third row by $-1/27$ gets us to the identity. Performing these five operations in succession to L^{-1} we see $\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -12 & 4 & 1 \end{bmatrix}$ go to $\begin{bmatrix} 1 & 0 & 0 \\ 3/27 & -1/27 & -7/27 \\ -12 & 4 & 1 \end{bmatrix}$, then $\begin{bmatrix} -9/27 & 12/27 & 3/27 \\ 3/27 & -1/27 & -7/27 \\ -12 & 4 & 1 \end{bmatrix}$, then $\begin{bmatrix} -6/27 & 11/27 & -4/27 \\ 3/27 & -1/27 & -7/27 \\ -12 & 4 & 1 \end{bmatrix}$, and then scaling rows two and three finally see $A^{-1} = \frac{1}{27}\begin{bmatrix} 6 & 11 & -4 \\ -3 & 1 & 7 \\ 12 & -4 & -1 \end{bmatrix}$.

3. Fundamental Subspaces

Let a $m \times n$ matrix A represent the linear transformation T between vector spaces V and W. A subset of a vector space that itself is a vector space is called a **subspace**. To check if a given space S is a subspace of V, we must verify that the set is closed under addition and scalar multiplication as defined in V.

The **Column Space** of the transformation is the set of all linear combination of the columns in the matrix. The system $Ax = b$ has solutions only when b lies in the column space of A; see that the column view of $Ax = b$ is $x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$. In this sense, the column space is the span of b. It is easy to see that this is a subspace. If two vectors, say x and x^*, lie in the column space, then $Ax = b$ and $Ax^* = b^*$. So $b + b^*$ must also be a combination of the columns, and surely $A(x + x^*) = (b + b^*)$. Also if any vector is a combination of the columns, then any multiple of the vector is also a combination of the columns; $Ax = b \Rightarrow A(cx) = cb$. We denote the column space $C(A)$.

The **Row Space** of the transformation is unsurprisingly the set of all linear combinations of the rows in a matrix. It is analogous to the column space of A^T since the columns in A^T are the rows in A. This is how we elect to refer to the subspace, and notion follows as such. We denote the column space of A^T $C(A^T)$.

The **Kernel or Null Space** of the transformation is $\ker(T) \coloneqq \{\vec{v} \in V | T(\vec{v}) = 0\}$. Geometrically, the null space represents the vectors of V which land on the origin after the linear transformation; they are the vectors x such that $Ax = 0$. or brevity, we can denote the null space of a linear transformation $N(A)$ from the matrix perspective. This is clearly a subspace; if $Ax = 0$ and $Ax^* = 0$, then $A(x + x^*)$. Likewise any multiple of zero is also zero, so scalar multiplication holds and we know the null space is a subspace of A. This is the reason why the vectors which fit a constraint form a subspace only when the constraint is zero—for any other vector, scalar multiplication or vector addition will break down. F

The **Right Null Space** of the transformation is the null space of A^T, and is denoted $N(A^T)$.

We now direct our attention to finding the dimension of each of the subspaces. The column space and null space are after different things—the column space is looking at the all the possible solution vectors b in the system $Ax = b$ while the null space is looking at the vectors x which yield the constraint of 0. Observe that the Column Space must be a subset of $\mathbb{R}^m$ since the scaled vectors adding to b are of length m. On the other hand, the Null Space must be a subset of

$\mathbb{R}^n$ since we are solving for the $n \times 1$ matrix which is multiplied by the $m \times n$ matrix to arrive at the $m \times 1$ matrix of zeros. The same operations applied to the transpose of the matrix will of course have reverse results. As $C(A) \subseteq \mathbb{R}^m$, $C(A^T) \subseteq \mathbb{R}^n$, and since $N(A) \subseteq \mathbb{R}^n$, $N(A^T) \subseteq \mathbb{R}^m$. Group the subspaces together to see $C(A) \cup N(A^T) \subseteq \mathbb{R}^m$ and $C(A^T) \cup N(A) \subseteq \mathbb{R}^n$. We call the dimension of the null space the **nullity**, and the dimension of the column space the **rank**. When the matrix is put into echelon form, the rank is the number of pivots in the matrix.

The first part of the Fundamental Theorem of Linear Algebra deals with the dimension of each of the subspaces. Where r is the number of pivots in the matrix, we can claim that $\dim(C(A) \subseteq \mathbb{R}^m) = r$, $\dim(N(A^T) \subseteq \mathbb{R}^m) = m - r$, $\dim(C(A^T) \subseteq \mathbb{R}^n) = r$, and $\dim(N(A) \subseteq \mathbb{R}^n) = n - r$.

Pf: We aim to show that the dimension of the kernel of T (called the nullity) plus the dimension of the image of T (called the rank) is equal to the dimension of the domain V.

Choose a basis for the null space, say $\langle \overrightarrow{v_1}, \dots, \overrightarrow{v_k} \rangle$. Then by adding additional vectors to this basis, one forms a basis for the entire Vector Space V. Call the basis which does this $\langle \overrightarrow{v_{k+1}}, \dots, \overrightarrow{v_n} \rangle$. Then the dimension of the Vector Space V is n. We have that the nullity is k, so it suffices to show that the rank is $r = n - k$ (and so the nullity is $k = n - r$).

Consider the vectors $\langle T(\overrightarrow{v_{k+1}}), \dots, T(\overrightarrow{v_n}) \rangle$ to be a basis for the column space. If this were the case, then the vectors would have to span the space. So choose an element $T(v)$ such that $v \in V$ (and then so that $T(v)$ is in the range space of the transformation). Since v is an element of V, it can be written as a linear combination of the basis we defined for V and constant terms $c_i \in \mathbb{R}$, namely $v = c_1 \cdot \overrightarrow{v_1} + \cdots + c_k \cdot \overrightarrow{v_k} + c_{k+1} \cdot \overrightarrow{v_{k+1}} + \cdots + c_n \cdot \overrightarrow{v_n}$. So one can call the element of the range space $T(v)$ the expanded function $T(c_1 \cdot \overrightarrow{v_1} + \cdots + c_k \cdot \overrightarrow{v_k} + c_{k+1} \cdot \overrightarrow{v_{k+1}} + \cdots + c_n \cdot \overrightarrow{v_n})$. Since linear transformations are structure preserving, we can break up the function like so $h(v) = c_1 L(\overrightarrow{v_1}) + \cdots + c_k L(\overrightarrow{v_k}) + c_{k+1} L(\overrightarrow{v_{k+1}}) + \cdots + c_n L(\overrightarrow{v_n})$.

Notice that $\langle \overrightarrow{v_1}, \dots, \overrightarrow{v_k} \rangle$ was previously defined to be the basis for the null space, the set of all elements in the domain which map to the 0 vector. So for all terms $T(\overrightarrow{v_1}) + \cdots + T(\overrightarrow{v_k})$, we have a value of 0, implying that $h(v)$ can be written $c_{k+1} T(\overrightarrow{v_{k+1}}) + \cdots + c_n T(\overrightarrow{v_n})$. In other words, the vectors $\overrightarrow{v_1}, \dots, \overrightarrow{v_k}$ are not relevant in the span. Counting terms, we have that the rank of the transformation is $n - k$ as desired.

We have shown that the basis $\langle T(\overrightarrow{v_{k+1}}), \dots, T(\overrightarrow{v_n}) \rangle$ spans the range space. We must verify that the choices of the vectors in the basis are independent. In doing so, we must show that no non-trivial choices for the scalars in the linear combination yield a value of 0. We assume the opposite, that there exists c_i's so $c_{k+1} T(\overrightarrow{v_{k+1}}) + \cdots + c_n T(\overrightarrow{v_n}) = 0$. Like before, we leverage the fact that the transformation is structure preserving to rewrite this combination of terms as

$T(c_{k+1}\overrightarrow{v_{k+1}} + \cdots + c_n\overrightarrow{v_n}) = 0$. Note that $c_{k+1}\overrightarrow{v_{k+1}} + \cdots + c_n\overrightarrow{v_n}$ must, by definition, be a member of the null space. Since $\langle\overrightarrow{v_1}, \ldots, \overrightarrow{v_k}\rangle$ is a basis for the null space, one can write $c_{k+1}\overrightarrow{v_{k+1}} + \cdots + c_n\overrightarrow{v_n}$ as a linear combination of scalars and $\langle\overrightarrow{v_1}, \ldots, \overrightarrow{v_k}\rangle$. We have that $c_{k+1}\overrightarrow{v_{k+1}} + \cdots + c_n\overrightarrow{v_n} = c_1 \cdot \overrightarrow{v_1} + \cdots + c_k \cdot \overrightarrow{v_k}$. By subtracting the left side from the right, we see that $(c_1 \cdot \overrightarrow{v_1} + \cdots + c_k \cdot \overrightarrow{v_k}) - (c_{k+1}\overrightarrow{v_{k+1}} + \cdots + c_n\overrightarrow{v_n}) = 0$. But $\langle\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}\rangle$ was defined to be a basis for V, and such a collection of vectors only sum to zero when all scalars in the linear combination are 0. This is a contradiction to our assumption, so verifies independence. Analogous results hold for the transposes of the matrix. **Q.E.D.**

The next part of the Fundamental Theorem of Linear Algebra deals with the direction of the subspaces. We've seen before that vectors x and y are orthogonal if their dot product $x^T y$ is zero. We call two subspaces orthogonal if every vector in one space is orthogonal to every vector in the other. Recall that the null space is defined to be all the solutions to $Ax = 0$. Multiplying this out, we see $\begin{bmatrix} \cdots & row\ 1 & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & row\ m & \cdots \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0_1 \\ \vdots \\ 0_m \end{bmatrix}$. The first component of the zero vector is a linear combination of the first row of A and the column matrix x. Likewise, the inner product of the second row of A and x is 0. So x is orthogonal to every row in A; $N(A) \perp C(A^T)$. We see the same result for the transpose, $N(A^T) \perp C(A)$. In a roundabout way, we see the problem $Ax = b$ in a different light. We have already shown how b must be a member of the column space to have a solution, and now we can say that $Ax = b$ has a solution only when b is orthogonal to every vector in the left null space.