Row Operations and Reduced Row Echelon Form

We now change gears and enter into a computation phase in the course. Without a doubt, a fundamental attribute of Linear Algebra is that it offers a number of rather straightforward techniques used in computations. We'll discuss a primary example of this in this section.

Suppose $A = [a_{ij}]$ is a $m \times n$ matrix. There are three elementary row operations which can be applied to A.

Operation I: We can multiply a row of A by a nonzero constant.

Operation II: We can add a scalar multiple of the j-th row of A to the i-th row of A, where $i \neq j$.

Operation III: We can interchange two distinct rows of A.

Proposition 16: Suppose A is a $m \times n$ matrix and B is a $m \times n$ matrix which is obtained by performing an elementary row operation to A. Then A can be obtained from B by performing an elementary row operation on B.

Moreover, the type of row operation performed on B to obtain A is the same as the type of row operation used to transform A into B.

Proof in the case B is obtained by performing a type I row operation on A: Suppose B is obtained from A by multiplying the i-th row of A by r, where $r \neq 0$. Then A is obtained from B by multiplying the i-th row of B by $\frac{1}{r}$.

HW #20: Complete the proof of

Proposition 16.

<u>Definition</u>: Consider the following situation:

$$A = A_0 \to A_1 \to \cdots \to A_k = B$$

where each A_i is a $m \times n$ matrix, for $i = 0, 1, \cdots, k$, and A_i is obtained from A_{i-1} by performing a single elementary row operation on A_{i-1}, for $i = 1, \cdots, k$. Then we say B is row equivalent to A.

Note that Proposition 16 implies

that if B is row equivalent to A, then A is row equivalent to B.

The elementary row operations are frequently employed to transform a matrix into a matrix which is in reduced row echelon form. We'll use RREF to denote reduced row echelon form. It is useful to introduce some terminology prior to defining the phrase "the $m \times n$ matrix B is in RREF".

Suppose $B = [b_{ij}]$ is a $m \times n$

matrix. If R_i denotes the i-th row of B, then

$$R_i = [b_{i1} \; b_{i2} \cdots b_{in}].$$

R_i is a zero row or row of zeros if $b_{ij} = 0$ for $j = 1, 2, \cdots, n$.

R_i is a nonzero row if there is at least one $k \in \{1, 2, \cdots, n\}$ such that $b_{ik} \neq 0$.

If R_i is a nonzero row, then the first nonzero term or entry in R_i is $b_{i\ell}$ - where $b_{i\ell} \neq 0$ but $b_{ij} = 0$ for $j = 1, \cdots, \ell - 1$. In words, if R_i is a nonzero row, then the first nonzero entry of R_i is just the first nonzero entry we encounter as we read R_i from left to right.

<u>Definition</u>: Let B be a $m \times n$ matrix whose i-th row is R_i, for $i = 1, \cdots, m$. We say that B is in reduced row echelon form (RREF) provided B satisfies the following four properties:

1) If R_i is a nonzero row of B then the first nonzero entry in R_i must be a 1. This

particular 1 is called a pivot, or to be more precise, the pivot in R_i.

2) If R_s and R_t are two nonzero rows of B, where $s<t$, if the pivot in R_s occurs in the j_s column of B and the pivot in R_t occurs in the j_t column of B, then $j_s < j_t$. (This means that pivots move from left to right as we read the rows of B from top to bottom.)

3) If a column of B contains a pivot, then every entry of this column, other than the pivot's entry, must be a zero.

4) If R_s is a nonzero row of B and R_t is a zero row of B, then $s<t$. (This means all the nonzero rows are above all the zero rows.)

The matrix, below, is in RREF – with the pivots circled. In this example,

$$B = \begin{bmatrix} 0 & 0 & 1 & 2 & 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 2 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

note that the first nonzero entry in each nonzero row of B is a 1, the pivot, the pivots move from left to right as we read the rows of B from top to bottom, each column of B containing a pivot has a zero in every entry – except for the pivot's entry and, lastly, all the nonzero rows of B are above all the zero rows of B.

HW #21: Determine which of the following matrices are in RREF.

(i) $\begin{bmatrix} 0 & 0 & 1 & 2 \end{bmatrix}$

(ii) $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

(iii) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(v) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

(vi) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(vii) $\begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$

(viii) $\begin{bmatrix} 0 & 0 & 1 & 2 & 3 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Theorem 9

The proof of the following Theorem provides an algorithm by which a matrix can be transformed, via row operations, into a matrix in RREF.

Theorem 9: Suppose A is a $m \times n$ matrix. Then A is row equivalent to a matrix in reduced row echelon form.

Proof: If A is the zero matrix, i.e. $A = O_{m,n}$, then A is already in RREF.

We now assume that A is a nonzero

matrix and proceed by induction.

As we proceed, we will construct a sequence of $m \times n$ matrices

$$A = A_0 \to A_1 \to \cdots \to A_k$$

where either $A_i = A_{i-1}$ or A_i is obtained by performing a single elementary row operation on A_{i-1}, for $i = 1, \cdots, k$. For notation we'll write:

$$A_r = \left[a_{ij}^{(r)} \right].$$

In words, A_r is a $m \times n$ matrix whose i,j-th entry is $a_{ij}^{(r)}$. The i-th row of A_r will be denoted by $R_i^{(r)}$, i.e.

$$R_i^{(r)} = [\, a_{i1}^{(r)} \; a_{i2}^{(r)} \; \cdots \; a_{in}^{(r)} \,],$$

and the j-th column of A_r will be denoted by $C_j^{(r)}$, i.e.

$$C_j^{(r)} = \begin{bmatrix} a_{1j}^{(r)} \\ a_{2j}^{(r)} \\ \vdots \\ a_{mj}^{(r)} \end{bmatrix}.$$

From the standpoint of the induction process it is useful to view A_r as being constructed from two sub-matrices of A_r - a top matrix, denoted by T_r, and a bottom matrix, denoted by B_r. T_r will consist of the top ℓ rows of A_r and B_r will consist of the bottom $m-\ell$ rows of A_r — where ℓ will depend on where we are in the inductive process.

Thus, pictorially we have:

$$A_r = \begin{bmatrix} T_r \\ \hdashline B_r \end{bmatrix} \begin{matrix} \} \text{ top } \ell \text{ rows of } A_r \\ \} \text{ bottom } m-\ell \text{ rows of } A_r . \end{matrix}$$

T_r will be viewed as a $\ell \times n$ matrix and B_r will be viewed as a $(m-\ell) \times n$ matrix.

To be more precise, T_r will be viewed as a $\ell \times n$ matrix in RREF.

Each of the ℓ rows of T_r will contain a pivot. The pivot in the i-th row of T_r will appear in the j_i-th column of T_r (or A_r). Note that $j_1 < j_2 < \cdots < j_\ell$ – since T_r is in RREF. Although B_r need not be in RREF, B_r will be required to satisfy the following property: Each entry in the first j_ℓ columns of B_r must be a zero.

Observe that this constraint on B_r depends only on the last row of T_r. This dependency is depicted below.

$$
\begin{array}{cl}
\left[\begin{array}{ccccccc}
0 & \cdots & 0 & 1 & a^{(r)}_{\ell\, j_\ell+1} & \cdots & a^{(r)}_{\ell n} \\
\hdashline
0 & \cdots & 0 & 0 & a^{(r)}_{\ell+1\, j_\ell+1} & \cdots & a^{(r)}_{\ell+1\, n} \\
\vdots & & \vdots & \vdots & \vdots & & \\
0 & \cdots & 0 & 0 & a^{(r)}_{m\, j_\ell+1} & \cdots & a^{(r)}_{mn}
\end{array}\right]
&
\begin{array}{l}
\leftarrow \text{last row of } T_r \\
\left.\begin{array}{l} \\ \\ \\ \end{array}\right\} B_r
\end{array}
\\
\uparrow & \\
j_\ell\text{-th column} &
\end{array}
$$

Base Case: We start with $A = A_0$ — knowing only that A_0 has a nonzero entry. Our goal is to show that A_0 is row equivalent to some $m \times n$ matrix A_s, where, basically, A_s gets the induction process started. By this we mean the first row of A_s serves as T_s while the last m-1 rows of A_s serve as B_s – where T_s and B_s have the properties discussed above.

To this end, let j_1 denote the smallest value of j such that the j-th column of A_0 has a nonzero entry. Put another way, the j_1-st column of A_0 has a nonzero entry and no column in A_0 to the left of the j_1-st column has a nonzero entry. We now focus on the j_1-st column of A_0, $C^{(0)}_{\cdot j_1}$, as depicted below.

$$C^{(0)}_{\cdot j_1} = \begin{bmatrix} a^{(0)}_{1j_1} \\ a^{(0)}_{2j_1} \\ \vdots \\ a^{(0)}_{mj_1} \end{bmatrix}$$

Let i_1 denote the smallest value of i such that $a^{(0)}_{ij_1} \neq 0$. In other words, as we read $C^{(0)}_{\cdot j_1}$, from top to bottom, the first nonzero entry we encounter is $a^{(0)}_{i_1 j_1}$.

If $i_1 = 1$, set $A_1 = A_0$. If $i_1 > 1$, let A_1 be the $m \times n$ matrix obtained by interchanging the 1-st and i_1-st rows of A_0. Note that in either case, every column to the left of the j_1-st column of A_1 has only zeros for entries and the $1, j_1$-entry of A_1, $a^{(1)}_{1 j_1}$, is nonzero.

If $a^{(1)}_{1 j_1} = 1$, set $A_2 = A_1$. If $a^{(1)}_{1 j_1} \neq 1$, let A_2 be the

$m \times n$ matrix obtained by multiplying the 1-st row of A_1 by $1/a^{(1)}_{1j_1}$. This is possible since $a^{(1)}_{1j_1} \neq 0$.

In either case, every column of A_2 to the left of its j_1-st column has only zeros for entries and $a^{(2)}_{1j_1} = 1$. As such, the first row of A_2, when viewed as a $1 \times n$ matrix, is in RREF and has a pivot in its j_1-st column. However, the last $m-1$ rows of A_2 may require more work.

Our present situation is as follows: The j_1-st column of A_2 is as depicted below.

$$C^{(2)}_{j_1} = \begin{bmatrix} 1 \\ a^{(2)}_{2j_1} \\ \vdots \\ a^{(2)}_{mj_1} \end{bmatrix}$$

In order to complete the Base Case, we must replace any nonzero $a^{(2)}_{ij_1}$ with a 0, for $i=2,\dots,m$.

If $a^{(2)}_{2j_1}=0$, let $A_3=A_2$. If $a^{(2)}_{2j_1}\neq 0$, let A_3 denote the $m\times n$

matrix obtained by adding $-a_{2j_1}^{(2)}$ times the 1-st row of A_2 to the second row of A_2.

In either case, every column to the left of the j_1-st column of A_3 has only zeros for entries and the j_1-st column of A_3 appears as below.

$$C_{j_1}^{(3)} = \begin{bmatrix} 1 \\ 0 \\ a_{3j_1}^{(3)} \\ \vdots \\ a_{mj_1}^{(3)} \end{bmatrix}$$

This last step provides us with a method for cleaning up the remainder of the j_1-st column. After $m-2$ analogous steps we arrive at A_s, where $s = 3+m-2 = m+1$, and A_s has the following properties: The entries in any column to the left of the j_1-st column of are zeros and the j_1-st column of A_s is as depicted below.

$$C_{j_1}^{(s)} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

As such, letting T_s denote the 1-st row of A_s and B_s denote the last $m-1$ rows of A_s - it follows that T_s is a $1 \times n$ matrix in RREF and having a pivot in its j_1-st column while the first j_1 columns of B_s have only zeros for entries.

This completes the Base Case.

<u>Inductive Step:</u> We now suppose that A is row equivalent to A_r where the first ℓ rows of A_r form T_r and the last $m-\ell$ rows of A_r form B_r.

Viewed as a $\ell \times n$ matrix, T_r is in RREF and the i-th row of T_r has a pivot in its j_i-th column, for $i = 1, \cdots, \ell$. Also, each entry in the first j_ℓ columns of B_r is a zero.

Note that if every entry in B_r is zero, then A_r is in RREF – and we are done.

If B_r has a nonzero entry, we must show that A_r is row equivalent to some $m \times n$ matrix A_t which enjoys the following properties:

1) If T_t denotes the first $\ell+1$ rows of A_t, then T_t is a $(\ell+1)\times n$ matrix in RREF and the i-th row of T_t has a pivot in its j_i-th column, for $i=1,\cdots,\ell+1$.

2) If B_t denotes the last $m-(\ell+1)$ rows of A_t, then every entry in the first $j_{\ell+1}$ columns of B_t is a zero.

Fortunately, the proof of the Inductive Step is very similar to the proof of the Base Case. In spirit, the initial phase of the Inductive Step's proof is to apply the steps used in the Base Case to B_r — rather than $A = A_0$. This transforms B_r into a $(m-\ell) \times n$ matrix, $B_{\tilde{s}}$ say, where the first row of $B_{\tilde{s}}$ is in RREF and whose last $m-\ell-1$ rows will eventually become B_t. When the top row of $B_{\tilde{s}}$ is added to the bottom of T_r, so as to form a $(\ell+1) \times n$ matrix, we may have to do some additional clean up work to obtain T_t.

We now enter into the details of the Inductive Step's proof. Let $j_{\ell+1}$ denote the smallest value of j such that the j-th column of B_r has a nonzero entry. Note that since the entries in the first j_ℓ columns of B_r all equal zero, $j_{\ell+1} > j_\ell$. The $j_{\ell+1}$-st column of A_r is as depicted below

$$C^{(r)}_{j_{\ell+1}} = \begin{bmatrix} a^{(r)}_{1 j_{\ell+1}} \\ \vdots \\ a^{(r)}_{\ell j_{\ell+1}} \\ \hdashline a^{(r)}_{\ell+1\, j_{\ell+1}} \\ \vdots \\ a^{(r)}_{m j_{\ell+1}} \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ \vdots \\ a \end{matrix}}\right\} \text{column in } T_r \\ \left.\vphantom{\begin{matrix} a \\ \vdots \\ a \end{matrix}}\right\} \text{column in } B_r \end{matrix}$$

Let $i_{\ell+1}$ denote the smallest value of i in the set $\{\ell+1,\cdots,m\}$ such that $a^{(r)}_{ij_{\ell+1}} \neq 0$. Note that $i_{\ell+1} \in \{\ell+1,\cdots,m\}$, thus $a^{(r)}_{i_{\ell+1}j_{\ell+1}}$ will be the first nonzero entry we encounter as we read the $j_{\ell+1}$-st column of B_r from top to bottom.

If $i_{\ell+1} = \ell+1$, set $A_{r+1} = A_r$. If $i_{\ell+1} > \ell+1$, let A_{r+1} denote the $m\times n$ matrix obtained by interchanging the $\ell+1$-st and $i_{\ell+1}$-st rows in A_r.

In either case, the first ℓ rows of

A_{r+1} are just the rows in T_r, the last $m-\ell$ rows in A_{r+1} have zeros for entries to the left of the $j_{\ell+1}$-st column and the $\ell+1$-st row of A_{r+1} has a nonzero entry, $a^{(r+1)}_{\ell+1\, j_{\ell+1}}$, in its $j_{\ell+1}$-st column.

If $a^{(r+1)}_{\ell+1\, j_{\ell+1}} = 1$, set $A_{r+2} = A_{r+1}$. If $a^{(r+1)}_{\ell+1\, j_{\ell+1}} \neq 1$, let A_{r+2} denote the $m \times n$ matrix obtained by multiplying the $\ell+1$-st row of A_{r+1} by $1/a^{(r+1)}_{\ell+1\, j_{\ell+1}}$. This is possible since $a^{(r+1)}_{\ell+1\, j_{\ell+1}} \neq 0$.

In either case, the first ℓ rows in A_{r+2} are just the rows in T_r, the last $m-\ell$ rows in A_{r+2} have zeros for entries to the left of the $j_{\ell+1}$-st column and $a^{(r+2)}_{\ell+1\, j_{\ell+1}} = 1$. Note that when viewed as a $1\times n$ matrix, the $\ell+1$-st row of A_{r+2} is in RREF and has a pivot in its $j_{\ell+1}$-st column.

We now focus on the $j_{\ell+1}$-st column of A_{r+2} – as depicted below.

$$C^{(r+2)}_{j_{\ell+1}} = \begin{bmatrix} a^{(r+2)}_{1 j_{\ell+1}} \\ \vdots \\ a^{(r+2)}_{\ell j_{\ell+1}} \\ 1 \\ a^{(r+2)}_{\ell+2\, j_{\ell+1}} \\ \vdots \\ a^{(r+2)}_{m j_{\ell+1}} \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \ell+1\text{-st row} \\ \\ \\ \\ \end{array}$$

The desired matrix to complete the Inductive Step, A_t, is obtained from A_{r+2} in $m-1$ steps. Each of these steps involve replacing $a_{ij_{\ell+1}}^{(r+2)}$ by zero, for $i=1,\cdots,\ell,\ell+2,\cdots,m$. Each of these steps is accomplished by adding an appropriate scalar multiple of the $\ell+1$-st row to the i-th row, for $i=1,\cdots,\ell,\ell+2,\cdots,m$.

This completes the proof of the Inductive Step - thus Theorem 9 is established.

Some Comments Concerning Theorem 9:

As noted prior to stating Theorem 9, the proof of Theorem 9 does provide an algorithm for transforming a $m \times n$ matrix into an $m \times n$ matrix in RREF. It treats an $m \times n$ matrix as a computer program would - so it may not be the most efficient approach for all matrices.

For instance, if you were to apply the steps in the proof of Theorem 9 to a matrix A, where A is already in RREF, A would remain unchanged throughout the entire process. This hints at the fact that every $m \times n$ matrix is row equivalent to a UNIQUE $m \times n$ matrix in RREF. This uniqueness statement will be addressed after we have discussed the row space of a matrix.