An Addendum To Theorem 9

Theorem 9 tells us that every $m \times n$ matrix is row equivalent to a matrix in RREF. Here we supplement this statement by showing that the matrix in RREF, provided by Theorem 9, is unique.

Proposition 18: Suppose A and B are two row equivalent $m \times n$ matrices and both A and B are in RREF. Then $A = B$.

Proof: Since A is row equivalent to

B, Proposition 16 implies

$$\text{row space of } A = \text{row space of } B.$$

This immediately implies that if either A or B is the zero matrix, then both must be the zero matrix.

We now assume that both A and B are nonzero matrices. Let

$$W = \text{row space of } A = \text{row space of } B.$$

Theorem 8 implies that any two bases for W must have the same number of elements. Since both A and B are in RREF, the nonzero rows of A forms a basis for W, as do the nonzero

rows of B - due to Proposition 17.
It follows that both A and B have the same number of nonzero rows, k say.

Let $\vec{v}_i$ denote the element of $\mathbb{R}^n$ corresponding to the i-th row of A, for $i = 1, \cdots, k$, and, similarly, let $\vec{w}_i$ denote the element of $\mathbb{R}^n$ corresponding to the i-th row of B, for $i = 1, \cdots, k$.
Also, let the columns of A containing a pivot be $j_1 < j_2 < \cdots < j_k$ and the columns of B containing a pivot be $\tilde{j}_1 < \tilde{j}_2 < \cdots < \tilde{j}_k$.

The key step in this proof is showing

$j_i = \tilde{j}_i$ for $i = 1, \cdots, k$. Since $j_1 < j_2 < \cdots < j_k$ and $\tilde{j}_1 < \tilde{j}_2 < \cdots < \tilde{j}_k$, if $\{j_1, j_2, \cdots, j_k\} = \{\tilde{j}_1, \tilde{j}_2, \cdots, \tilde{j}_k\}$, it would follow that $j_i = \tilde{j}_i$, for $i = 1, \cdots, k$.

Thus, we must show that

$$\{j_1, \cdots, j_k\} = \{\tilde{j}_1, \cdots, \tilde{j}_k\}.$$

To this end, suppose that

$$\{j_1, \cdots, j_k\} \neq \{\tilde{j}_1, \cdots, \tilde{j}_k\}.$$

It would follow that there is either a j_r such that $j_r \notin \{\tilde{j}_1, \cdots, \tilde{j}_k\}$ or there is a $\tilde{j}_s$ such that $\tilde{j}_s \notin \{j_1, \cdots, j_k\}$.

(Actually, both would have to be true, by a counting argument - but we don't need this.)

Without loss of generality, we will assume there is a $\tilde{j}_s$ such that $\tilde{j}_s \notin \{j_1, \cdots, j_k\}$. Since both $\{\vec{v}_1, \cdots, \vec{v}_k\}$ and $\{\vec{w}_1, \cdots, \vec{w}_k\}$ are bases for W, we must have $\vec{w}_s \in \text{span}(\{\vec{v}_1, \cdots, \vec{v}_k\})$. Thus, there exist $c_1, \cdots, c_k \in \mathbb{R}$ such that

$$\vec{w}_s = \sum_{i=1}^{k} c_i \vec{v}_i \,.$$

Recall, from the proof of Proposition 17, that the j_i-th component of $\sum_{i=1}^{k} c_i \vec{v}_i$ is c_i, for $i = 1, \cdots, k$. Since the first nonzero component of $\vec{w}_s$ is a 1 in its $\tilde{j}_s$-th component, $c_i = 0$ if $j_i < \tilde{j}_s$.

This immediately implies that j_i cannot be less than $\tilde{j}_s$ for all $i=1,\cdots,k$.

Consequently, there is a smallest t such that $\tilde{j}_s < j_t$. Since $c_i = 0$ if $i < t$, it follows that the first possible nonzero component of $\sum_{i=1}^{k} c_i \vec{v}_i$ is its j_t-th component. But $\vec{w}_s$ has a 1 in its $\tilde{j}_s$-th component and $\tilde{j}_s < j_t$. Thus, $\vec{w}_s \notin \operatorname{span}(\{\vec{v}_1,\cdots,\vec{v}_k\})$ – a contradiction. This contradiction implies that $\{j_1,\cdots,j_k\} = \{\tilde{j}_1,\cdots,\tilde{j}_k\}$ – so $j_i = \tilde{j}_i$, for $i=1,\cdots,k$.

We are now in a position to show that $\vec{v}_i = \vec{w}_i$, for $i=1,\cdots,k$. To see this, fix $\ell \in \{1, \cdots, k\}$. Note that we must have

$$\vec{w}_\ell \in \operatorname{span}(\{\vec{v}_1, \cdots, \vec{v}_k\}).$$

Thus, there exist $r_1, \cdots, r_k \in \mathbb{R}$ such that

$$\vec{w}_\ell = \sum_{i=1}^{k} r_i \vec{v}_i .$$

Now note that the j_i-th component of $\sum_{i=1}^{k} r_i \vec{v}_i$ is r_i, for $i = 1, \cdots, k$. On the other hand, since $\tilde{j}_i = j_i$, for $i = 1, \cdots, k$, the j_i-th component of $\vec{w}_\ell$ is 0, if $i \neq \ell$, and 1 if $i = \ell$. Thus, $r_i = 0$ if $i \neq \ell$. This implies that

$$\vec{w}_\ell = r_\ell \vec{v}_\ell .$$

Since both $\vec{w}_\ell$ and $\vec{v}_\ell$ have a 1 in their

j_ℓ-th component, we conclude that $r_\ell = 1$ — so $\vec{w}_\ell = \vec{v}_\ell$.

Since $\vec{v}_i = \vec{w}_i$, for $i = 1, \cdots, k$, it follows that the first k rows in A are the same as the first k rows in B. Since these are the only nonzero rows in A and B, and A and B are both $m \times n$ matrices — we must have that $A = B$.

<u>HW #23</u>: *Proposition 18 provides the last element needed to prove that there is a one-to-one correspondence between subspaces of $\mathbb{R}^n$ and the*

set of $n \times n$ matrices in RREF.

Argue that every 2×2 matrix in RREF is of one of the four types of matrices, listed below, and draw a picture in $\mathbb{R}^2$ of the subspace represented by each of these four types of matrices.

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

(ii) $\begin{bmatrix} 1 & m \\ 0 & 0 \end{bmatrix}$ $\quad m \in \mathbb{R}$

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

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

Systems of Linear Equations

Many students are exposed to Gaussian Elimination, and possibly matrices in RREF, when solving systems of linear equations in Precalculus courses — so you may find this section as being a review, for the most part. Recall that the solution space of a system of homogeneous, linear equations in n variables is a subspace of $\mathbb{R}^n$. We'll show how to construct a basis for such a nonzero solution space.

We'll employ the following notation to describe a generic system of m linear equations in n variables:

$$\mathcal{S}: \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= c_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= c_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= c_m \end{aligned} \quad .$$

– where $x_1, x_2, \ldots, x_n$ are the n variables, while the a_{ij}'s and c_i's represent fixed constants for all $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

You probably noticed that the notation used for the system $\mathcal{S}$, above, is reminiscent of matrix notation. This is no fluke – matrix techniques, and the notion of RREF, provide a very efficient approach to determining the solution set of a system of linear equations.

Recall that two systems of equations, $\mathcal{S}$ and $\tilde{\mathcal{S}}$ say, are said to be equivalent precisely when the solution set of $\mathcal{S}$ equals the solution set of $\tilde{\mathcal{S}}$. If $\mathcal{S}$ is a system of equations, there are three elementary operations which can be performed on $\mathcal{S}$ to yield an equivalent system of equations:

I) Multiply an equation of $\mathcal{S}$ by a nonzero real number.

II) Add a real number multiple of one equation in $\mathcal{S}$ to a different equation in $\mathcal{S}$.

III) Change the order in which the equations in $\mathcal{S}$ appear.

We now associate the generic system of linear equations, $\mathcal{S}$, on page 222, to the $m \times (n+1)$ matrix shown at the top of page 225. Such a matrix is called an augmented matrix.

$$A_{\mathcal{S}} = \left[\begin{array}{cccc:c} a_{11} & a_{12} & \cdots & a_{1n} & c_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & c_2 \\ \vdots & \vdots & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & c_m \end{array}\right].$$

Observe that the three elementary operations on $\mathcal{S}$, given on page 224, correspond to the three elementary row operations we can apply to $A_{\mathcal{S}}$. This leads to

<u>Proposition 19</u>: Suppose $\mathcal{S}$ and $\tilde{\mathcal{S}}$ are two systems of linear equations, each consisting of m equations in n variables. Let $A_{\mathcal{S}}$ and $A_{\tilde{\mathcal{S}}}$ denote the matrices associated to $\mathcal{S}$ and $\tilde{\mathcal{S}}$, respectively.

If $A_{\mathcal{S}}$ is row equivalent to $A_{\tilde{\mathcal{S}}}$, then $\mathcal{S}$ and $\tilde{\mathcal{S}}$ are equivalent systems of equations. I.e. the solution set of $\mathcal{S}$ equals the solution set of $\tilde{\mathcal{S}}$.

To implement Proposition 19 we proceed as follows:

<u>Step 1</u>: Given $\mathcal{S}$, as on page 222, construct $A_{\mathcal{S}}$, as on page 225.

<u>Step 2</u>: Transform $A_{\mathcal{S}}$ into a matrix, $A_{\tilde{\mathcal{S}}}$, in RREF.

<u>Step 3</u>: Construct $\tilde{\mathcal{S}}$ from $A_{\tilde{\mathcal{S}}}$ and solve $\tilde{\mathcal{S}}$.

This process is illustrated in the next example.

<u>Ex</u>: Solve the system

$$\mathcal{S}: \quad \begin{aligned} x + 2y + 3z &= 4 \\ 5x + 6y + 7z &= 8 \\ 9x + 10y + 11z &= 12 \end{aligned} .$$

$$A_{\mathcal{S}} = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 9 & 10 & 11 & 12 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -24 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & -8 & -16 & -24 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So $$A_{\tilde{S}} = \left[\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\tilde{S}: \quad \begin{aligned} x \qquad - z &= -2 \\ y + 2z &= 3 \\ 0 &= 0 \end{aligned}$$

$\tilde{S}$ can be solved by rewriting the equations in $\tilde{S}$. This involves moving those expressions involving the variables, which correspond to the columns of $A_{\tilde{S}}$ not containing a pivot, to the right-hand side of each equation.

So we write $\tilde{\mathcal{S}}$ as below:

$$\tilde{\mathcal{S}}: \quad \begin{aligned} x &= -2+z \\ y &= 3-2z \\ 0 &= 0 \end{aligned}$$

Now, according to Proposition 19, the solution set of $\mathcal{S}$ equals the solution set of $\tilde{\mathcal{S}}$

$$= \{ [-2+z, 3-2z, z] \mid z \in \mathbb{R} \}$$

The verification of this is given below:

$(-2+z) + 2(3-2z) + 3z = -2+6 = 4$ ✓

$5(-2+z) + 6(3-2z) + 7z = -10+18 = 8$ ✓

$9(-2+z) + 10(3-2z) + 11z = -18+30 = 12$ ✓

<u>Some Useful Terminology</u>: Note that when

$\mathcal{S}$ is associated to $A_{\mathcal{S}}$, the coefficients of the i-th variable become the entries is the i-th column of $A_{\mathcal{S}}$ and the constants become the entries in the augmentation column of $A_{\mathcal{S}}$. In this way, we view the i-th column of $A_{\mathcal{S}}$ as the column associated to x_i and the augmentation column of $A_{\mathcal{S}}$ as the constant column.

After transforming $A_{\mathcal{S}}$ into $A_{\tilde{\mathcal{S}}}$, the i-th column of $A_{\tilde{\mathcal{S}}}$ remains the column associated to x_i and the augmentation column of $A_{\tilde{\mathcal{S}}}$ corresponds to constants still.

As was illustrated in the last example, when $A_{\tilde{S}}$ is in RREF, the tried and true approach to constructing the solution set for $\tilde{S}$ is to move those expressions involving variables which are associated to columns of $A_{\tilde{S}}$ not involving a pivot to the right-hand side of each equation in $\tilde{S}$. When $\tilde{S}$ has a nonempty solution set, those variables associated to columns of $A_{\tilde{S}}$ containing a pivot should be viewed as dependent variables (like x and y in the last example) while those variables associated to columns of $A_{\tilde{S}}$

not containing a pivot should be viewed as independent variables (like z in the last example).

Suppose now that:

i) $\mathcal{S}$ is a system of m linear equations in the variables $x_1, \dots, x_n$,

ii) $A_{\mathcal{S}}$ is the matrix associated to $\mathcal{S}$,

iii) $A_{\mathcal{S}}$ is row equivalent to $A_{\tilde{\mathcal{S}}}$, where $A_{\tilde{\mathcal{S}}}$ is in RREF, and

iv) $\tilde{\mathcal{S}}$ is the system of m linear equations in the variables, $x_1, \dots, x_n$, associated to $A_{\tilde{\mathcal{S}}}$.

We now provide three basic observations

corresponding to the solution set of $\tilde{\mathcal{S}}$, hence the solution set of $\mathcal{S}$.

<u>Observation 1)</u>: If $A_{\tilde{\mathcal{S}}}$ has a pivot in its augmentation column, then $\mathcal{S}$ has no solutions.

<u>Reason</u>: This corresponds to the equation $0x_1 + \cdots + 0x_n = 1$, which has no solutions.

<u>Observation 2)</u>: If $A_{\tilde{\mathcal{S}}}$ has a pivot in every column, except the augmentation column, then $\mathcal{S}$ has a unique solution.

<u>Reason</u>: In this case, every nonzero equation of $\tilde{\mathcal{S}}$ is of the form

$$x_i = d_i$$

for some $d_i \in \mathbb{R}$, $i = 1, \cdots, n$. Thus

$\mathcal{S}$ has the unique solution

$[d_1, \cdots, d_n]$.

<u>Observation 3</u>: If $A_{\tilde{\mathcal{S}}}$ has at least two columns not containing a pivot, and one of these columns not containing a pivot is the augmentation column, then $\mathcal{S}$ has infinitely many solutions.

<u>Reason</u>: The solution set is not empty and one component in your general solution is an independent variable -

independent variables are allowed to be any real number.

In the case when $\mathcal{S}$ is also assumed to be homogeneous, more can be said.

<u>Observation 4</u>: If $\mathcal{S}$ is a system of m homogeneous linear equations in the variables $x_1, \ldots, x_n$, then the zero vector in $\mathbb{R}^n$ is a solution to $\mathcal{S}$.

<u>Reason</u>: If $\mathcal{S}$ is homogeneous, then

$\mathcal{S}$ is of the form depicted below.

$$\mathcal{S}: \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{array}$$

Clearly, the zero vector in $\mathbb{R}^n$ is a solution to such a system.

Actually, more can be said when $\mathcal{S}$ is homogeneous. As we noted earlier in this section, the solution space of $\mathcal{S}$ is a subspace of $\mathbb{R}^n$. As the next example suggests, when $\mathcal{S}$ is homogeneous, the dimension of the solution space of $\mathcal{S}$ equals the number of independent variables in its generic solution.

Ex: Solve the system $\mathcal{S}$, given below, and construct a basis for its solution space - viewed as a subspace of $\mathbb{R}^n$.

$$\mathcal{S}: \quad \begin{array}{rrrrrl} x_1 & & & -4x_4 & +5x_5 & =0 \\ & x_2 & +3x_3 & & -6x_5 & =0 \end{array}$$

In this case

$$A_8 = \left[\begin{array}{ccccc|c} 1 & 0 & 0 & -4 & 5 & 0 \\ 0 & 1 & 3 & 0 & -6 & 0 \end{array}\right].$$

A_8, as given, is in RREF. So to solve the system $\mathcal{S}$, we rewrite $\mathcal{S}$ as below.

$$\mathcal{S}: \begin{aligned} x_1 &= \; 4x_4 - 5x_5 \\ x_2 &= -3x_3 \qquad\; + 6x_5 \end{aligned}$$

The solution space for $\mathcal{S}$ is

$$\{ [4x_4 - 5x_5, -3x_3 + 6x_5, x_3, x_4, x_5] \mid x_3, x_4, x_5 \in \mathbb{R} \}.$$

To obtain a basis for the solution space of $\mathcal{S}$, we proceed as follows:

The first basis element is obtained by setting $x_3 = 1, x_4 = 0, x_5 = 0$.

We obtain $[0, -3, 1, 0, 0]$.

The second basis element is obtained by setting $x_3 = 0, x_4 = 1, x_5 = 0$.

We obtain $[4, 0, 0, 1, 0]$.

The third basis element is obtained by setting $x_3 = 0, x_4 = 0, x_5 = 1$.

We obtain $[-5, 6, 0, 0, 1]$.

We claim that

$$\{[0,-3,1,0,0], [4,0,0,1,0], [-5,6,0,0,1]\}$$

is a basis for this solution space.

Indeed, the generic element in the span of this set of vectors is

$$x_3[0,-3,1,0,0]+x_4[4,0,0,1,0]+x_5[-5,6,0,0,1]$$

$$=[4x_4-5x_5,-3x_3+6x_5,x_3,x_4,x_5]$$

-which is the generic solution to $\mathcal{S}$.

Thus, these vectors span the solution space of $\mathcal{S}$. Also,

$$[4x_4-5x_5,-3x_3+6x_5,x_3,x_4,x_5]$$

$$=[0,0,0,0,0]$$

implies $x_3=x_4=x_5=0$

-so these three vectors form a linearly independent set.

<u>HW #24</u>: In the example discussed on page 208, we showed that

$[1,1,1] \in \operatorname{span}(\{[1,2,3],[4,5,6],[7,8,9]\})$.

Verify this using the techniques of this section.

<u>Hint</u>: The typical element in this span can be written as

$$x[1,2,3]+y[4,5,6]+z[7,8,9].$$

You must show that the system of equations determined by

$$x[1,2,3]+y[4,5,6]+z[7,8,9]=[1,1,1]$$

has a solution.

Inverse Matrices - An Introduction

Recall that I_n denotes the $n \times n$ identity matrix.

We say that the $n \times n$ matrix A is invertible (or nonsingular) provided there is an $n \times n$ matrix B such that

$$AB = I_n = BA.$$

In this case, we'll frequently write $B = A^{-1}$.

Note that $I_n I_n = I_n$ - so I_n is invertible an $I_n^{-1} = I_n$. Since

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is invertible, and

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Also

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

So $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is invertible, and

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

The following Proposition provides two basic properties of inverse matrices and invertible matrices.

Proposition 20:

1) If A is invertible, then its inverse is unique.

2) If A and B are two invertible $n\times n$ matrices, then AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof:

1) Suppose A is an $n\times n$ invertible matrix and C and D are two $n\times n$ matrices such that $AC = I_n = CA$ and $AD = I_n = DA$.

Then

$$C = CI_n = C(AD) = (CA)D = I_n D = D.$$

2) This follows from

$$B^{-1}A^{-1}AB = B^{-1}I_nB = B^{-1}B = I_n$$

and

$$ABB^{-1}A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

<u>HW #25</u>: Suppose m is an integer and $m \geq 2$. Suppose further that $A_1, \cdots, A_m$ are $n \times n$ invertible matrices. Give an inductive proof that $A_1 \cdots A_m$ is invertible and $(A_1 \cdots A_m)^{-1} = A_m^{-1} \cdots A_1^{-1}$.

<u>Hint</u>: The Base Case is a consequence of Proposition 20, property 2).

It isn't difficult to construct noninvertible $n \times n$ matrices, consider $O_{n,n}$ for instance. The following Proposition is useful in this regard.

Proposition 21:

1) Suppose A is a $n \times n$ matrix whose k-th row is a row of zeros. Then A is not invertible.

2) If B is a $n \times n$ matrix in RREF, then B is invertible if and only if $B = I_n$.

Proof:

1) If B is any $n \times n$ matrix and

the k-th row of A consists solely of zeros, then the k-th row of AB consists solely of zeros. As such $AB \neq I_n$.

2) If $B = I_n$, then B is invertible.

Suppose now that B is invertible and B is in RREF. According to 1), B cannot have a row of zeros. This implies that B has a pivot in each of its n rows. This, in turn, implies that B has a pivot in each of its n columns. This implies that $B = I_n$.

Elementary Matrices

Suppose E is a $n \times n$ matrix which is obtained by performing a single elementary row operation on I_n. Then we say that E is an elementary matrix.

If E is an elementary $n \times n$ matrix and A is a $n \times k$ matrix, then EA is a $n \times k$ matrix which can be described as follows: EA is obtained from A by the same elementary row operation that is performed on I_n to obtain E.

We'll employ the following HW problem to

HW #26: Let $r, s \in \{1, \cdots, n\}$ and define M_{rs} to be the $n \times n$ matrix whose only nonzero entry is a 1 in the r,s-entry. I.e. $M_{rs} = [m_{ij}]$ where

$$m_{ij} = \begin{cases} 0 \text{ if either } i \neq r \text{ or } j \neq s \\ 1 \text{ if } i = r \text{ and } j = s \end{cases}.$$

Also, let A denote a $n \times k$ matrix. Then $M_{rs} A$ is the $n \times k$ matrix whose only nonzero row (potentially) is its r-th row and the r-th row of $M_{rs} A$ is the s-th row of A.

Hint: Recall that the i,j-th entry of $M_{rs} A$ is $\sum_{t=1}^{n} m_{it} a_{tj}$ where

$A = [a_{ij}]$.

Now, to establish the previously mentioned property of EA, we'll consider the elementary row operations separately.

<u>Case I</u>: Let r denote a nonzero real number and fix $\ell \in \{1, \cdots, n\}$. Suppose E is obtained by multiplying the ℓ-th row of I_n by r. Then EA is obtained from A by multiplying the ℓ-th row of A by r.

<u>Proof</u>: First note that $E = I_n - M_{\ell\ell} + rM_{\ell\ell}$.

Thus $EA = I_n A - M_{\ell\ell} A + r(M_{\ell\ell} A)$.

$I_n A = A$ together with HW #26 imply that $I_n A - M_{\ell\ell} A$ is the $n \times k$ matrix whose i-th row is the i-th row of A, for $i \neq \ell$, and whose ℓ-th row is a row of zeros.

Adding $r(M_{\ell\ell} A)$ to $I_n A - M_{\ell\ell} A$ replaces the row of zeros with r times the ℓ-th row of A. Thus EA is obtained from A by multiplying the ℓ-th row of A by r.

<u>Case II</u>: Suppose $c \in \mathbb{R}$, $s, t \in \{1, \cdots, n\}$ and $s \neq t$. Suppose E is obtained from I_n by adding c times the

t-th row of I_n to the s-th row of I_n. Then EA is obtained by adding c times the t-th row of A to the s-th row of A.

<u>Proof</u>: Note that $E = I_n + c\,M_{st}$. The result now follows immediately from $I_n A = A$ and $M_{st} A$ being the $n \times k$ matrix whose only possible nonzero row is its s-th row and the s-th row of $M_{st} A$ is the t-th row of A.

<u>Case III</u>: Let $s, t \in \{1, \cdots, n\}$, where $s \neq t$. Let E be obtained by interchanging the s-th and t-th rows of I_n.

Then EA is obtained by interchanging the s-th and t-th rows of A.

<u>Proof</u>: Observe that $E = I_n - M_{ss} - M_{tt} + M_{st} + M_{ts}$. As suggested by the proof of Case I, $I_n A - M_{ss}A - M_{tt}A$ is the $n \times k$ matrix whose i-th row is the i-th row of A, when $i \neq s, t$, and the s-th and t-th rows are rows of zeros. Adding $M_{st}A$ and $M_{ts}A$ to $I_n A - M_{ss}A - M_{tt}A$ replaces the zeros in row s by the t-th row of A and the zeros in row t by the s-th row of A.

HW #27: Show that every elementary matrix is invertible.

Hint: See Proposition 16, page 144.

Invertible Matrices and Row Operations

An immediate consequence of the last section is that if B is a $n \times k$ matrix obtained by performing a single elementary row operation on the $n \times k$ matrix A, then there is a $n \times n$ elementary matrix E such that $EA = B$.

Suppose now that B is row equivalent to A. Then there exist $n \times k$ matrices $A_0, A_1, \cdots, A_m$ such that

(i) $A = A_0$ and $B = A_m$,

(ii) A_i is obtained by performing a

single elementary row operation on A_{i-1}, for $i=1,\cdots,m$.

(ii) implies there are elementary $n\times n$ matrices $E_1,\cdots,E_m$ such that

$$E_i A_{i-1} = A_i$$

for $i=1,\cdots,m$.

Note then that $E_1 A_0 = A_1$ - so

$A_2 = E_2 A_1 = E_2 E_1 A_0$ - so

$A_3 = E_3 A_2 = E_3 E_2 E_1 A_0$ - so

$A_4 = E_4 A_3 = E_4 E_3 E_2 E_1 A_0$.

Proceeding in this manner we eventually obtain $A_m = E_m E_{m-1} \cdots E_2 E_1 A_0$

or $B = E_m \cdots E_1 A$. According to

a combination of HW #25 and HW #27, $B = E_m \cdots E_1 A$ can be re-written as $A = E_1^{-1} \cdots E_m^{-1} B$.

<u>Theorem 10</u>: Suppose A is a $n \times n$ matrix which is row equivalent to the $n \times n$ matrix B, where B is in RREF. Then A is invertible if and only if $B = I_n$.

<u>Proof</u>: Since A is row equivalent to B, there exist $n \times n$ elementary matrices $E_1, \cdots, E_m$ such that

$$E_m \cdots E_1 A = B.$$

HW #25 and HW #27 imply that,

if A is invertible, then $B = E_m \cdots E_1 A$ must also be invertible. Since B is in RREF, B invertible implies that $B = I_n$ by Proposition 21.

On the other hand, if $B = I_n$, then $A = E_1^{-1} \cdots E_m^{-1} B = E_1^{-1} \cdots E_m^{-1} I_n$ - so A is invertible. This completes the proof of Theorem 10.

Suppose now that A is an invertible $n \times n$ matrix. Theorem 10 implies that there exist elementary $n \times n$ matrices $E_1, \cdots, E_m$ such that $E_m \cdots E_1 A = I_n$. This implies $A = E_1^{-1} \cdots E_m^{-1} I_n = I_n E_1^{-1} \cdots E_m^{-1}$.

This, in turn, implies that $AE_m \cdots E_1 = I_n$.

Note that $E_m \cdots E_1 A = I_n = A E_m \cdots E_1$

– so $A^{-1} = E_m \cdots E_1$.

Now consider the $n \times (2n)$ matrix

$$\left[\begin{array}{c|c} A & I_n \end{array}\right].$$

Observe that

$$E_m \cdots E_1 \left[\begin{array}{c|c} A & I_n \end{array}\right]$$

$$= E_m \cdots E_1 \left(\left[\begin{array}{c|c} A & O_{nn} \end{array}\right] + \left[\begin{array}{c|c} O_{nn} & I_n \end{array}\right] \right)$$

$$= \left[\begin{array}{c|c} E_m \cdots E_1 & O_{nn} \end{array}\right] + \left[\begin{array}{c|c} O_{nn} & E_m \cdots E_1 I_n \end{array}\right]$$

$$= \left[\begin{array}{c|c} I_n & O_{nn} \end{array}\right] + \left[\begin{array}{c|c} O_{nn} & A^{-1} \end{array}\right]$$

$$= \left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right].$$

Since $E_m \cdots E_1$ corresponds to the sequence of elementary row operations which transform A into I_n, we now have a rather straightforward way to construct A^{-1} – assuming A is invertible. Begin with the matrix

$$\left[\begin{array}{c|c} A & I_n \end{array}\right].$$

We perform row operations on

$$\left[\begin{array}{c|c} A & I_n \end{array}\right]$$

which transform the left-hand side, i.e. A, into I_n. These row operations with then transform the right-hand side, I_n, into A^{-1}.

<u>Ex</u>: Compute A^{-1} if

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array}\right]$$

so $A^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

Check:

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \checkmark$$

Actually, the check above employed more work than was needed.

<u>Proposition 22</u>: Suppose A and B are $n \times n$ matrices such that

$AB = I_n$. Then A and B are both invertible and $B^{-1} = A$, $B = A^{-1}$.

<u>Proof</u>: Suppose A is row equivalent to C, where C is in RREF. Let $E_1, \cdots, E_m$ denote elementary matrices such that

$$E_m \cdots E_1 A = C.$$

Note that $AB = I_n$ implies

$$E_m \cdots E_1 AB = E_m \cdots E_1 I_n$$

or

$$CB = E_m \cdots E_1.$$

Now, if C has a row of zeros, then

CB has a row of zeros. However, $E_m \cdots E_1$ is invertible – so $E_m \cdots E_1$ has no row of zeros. As such, C cannot have a row of zeros. Since C is an $n \times n$ matrix in RREF and C has no row of zeros, $C = I_n$. This implies A is row equivalent to I_n – so A is invertible and $A^{-1} = E_m \cdots E_1$.

Also $C = I_n$ implies

$B = I_n B = CB = E_m \cdots E_1$ – so B is invertible and $B = A^{-1}$.

<u>HW #28</u>: Determine A^{-1} if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}.$$