Elementary Matrices

In this special handout of material not contained in the text, we introduce the concept of elementary matrix. Elementary matrices are useful in several ways that will be shown in this handout. One important use that we will examine first is that elementary matrices can be used to carry out elementary row operations. This means that if you want to interchange two rows, or multiply a row by a constant and add it to another row, or multiply some row by a non-zero constant, it can be done using an elementary matrix. The definition given next shows that elementary matrices can be created by applying elementary row operations to the identity matrix.

Definition: An nxn matrix E is an elementary matrix if it can be obtained by performing a single elementary row operation on the identity matrix I_n.

Some examples of elementary matrices for n = 3 and for each of the elementary row operations are

$$E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$, which is obtained by interchanging the first row and third rows of I_3,

$$E_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ which is obtained by multiplying the first row of I_3 by -1, and

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ which is obtained by multiplying the first row of I_3 by -2 and adding it to the second row of I_3. The subscripts on E have no particular meaning but are just used to distinguish one elementary matrix from the next.

An elementary row operation can be carried out by its corresponding elementary matrix through matrix multiplication. For example, suppose we wish to interchange the first and third rows of the matrix $A = \begin{bmatrix} 0 & 1 & -1 \\ 2 & -6 & 2 \\ -1 & 3 & 0 \end{bmatrix}$. This can be done by multiplying A on the left by the elementary

matrix E_1, given above, to yield

$$E_1 A = \begin{bmatrix} -1 & 3 & 0 \\ 2 & -6 & 2 \\ 0 & 1 & -1 \end{bmatrix}.$$

Suppose we then wish to perform the elementary row operation of multiplying the first row by -1, we can do this by multiplying by E_2 which then gives

$$E_2 E_1 A = \begin{bmatrix} 1 & -3 & 0 \\ 2 & -6 & 2 \\ 0 & 1 & -1 \end{bmatrix}$$

We can then perform on this result the elementary row operation of multiplying the first row by -2 and adding it to the second row using E_3 to obtain

$$E_3 E_2 E_1 A = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix}$$

If our goal is to obtain rref for A, then we would continue with the following elementary matrices to

switch rows 2 and 3 we do this to I_3 to obtain $E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ and applying this to the

previous result we get

$$E_4 \; E_3 \; E_2 \; E_1 \; A = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}$$

then to multiply row 2 by 3 and add it to row 1 we do this to I_3 to obtain $E_5 = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

and applying this to the previous result we obtain

$$E_5 E_4 E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}$$

then to multiply row 3 by 1/2 we use the elementary matrix $E_6 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$ to get

$$E_6 E_5 E_4 E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

then to multiply row 3 by 3 and add it to row 1 we use $E_7 = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ to obtain

$$E_7 E_6 E_5 E_4 E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

and finally, to add row 3 to row 2 we use $E_8 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ to get

$$E_8 E_7 E_6 E_5 E_4 E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that the last equation is $E_8 E_7 E_6 E_5 E_4 E_3 E_2 E_1 A = I_3$. Thus the product $E_8 E_7 E_6 E_5 E_4 E_3 E_2 E_1$ must be the inverse of A and this gives us another way of computing the inverse of A. Note that carrying out the product

$$E_8 E_7 E_6 E_5 E_4 E_3 E_2 E_1 = \begin{bmatrix} 3 & 3/2 & 2 \\ 1 & 1/2 & 1 \\ 0 & 1/2 & 1 \end{bmatrix} = A^{-1}$$

Elementary matrices have another important property given in the next theorem.

Theorem 1. Every elementary matrix has an inverse, which is also an elementary matrix.

Proof: We won't give a formal proof but will suggest two reasons why it is true. First, since every elementary row operation can be reversed, one would expect that this can be represented by an elementary row operation which is an elementary matrix. Second, it is easy to construct the inverse of an elementary matrix. Here are the three cases:

1. If E interchanges two rows, then E is its own inverse. (Why?)

2. If E multiplies a row by a non-zero constant c, then its inverse is the elementary matrix that multiplies a row by $1/c$.

3. If E multiplies row i by c and adds it to row j, the its inverse is the elementary matrix that multiplies row i by -c and adds it to row j.

For the elementary matrices used in the example above here are their inverses which are also elementary matrices

$$(E_1)^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad (E_2)^{-1} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (E_3)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(E_4)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad (E_5)^{-1} = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (E_6)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$(E_7)^{-1} = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (E_8)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Theorem 2. (a). Every mxn matrix A can be decomposed as a product of elementary matrices and a matrix R that is rref, that is,

$$A = \tilde{E}_1 \ldots \tilde{E}_{s-1} \; \tilde{E}_s R \tag{1}$$

where $\tilde{E}_i$ are elementary matrics and R is rref.

(b). If A is nxn and has an inverse, then R is the identity matrix and A is decomposable into the product of elementary matrices, that is,

$$A = \tilde{E}_1 \ldots \tilde{E}_{s-1} \; \tilde{E}_s \tag{2}$$

Proof of (a). Since every matrix A can be put into rref by row operations and since every row operation can be represented by an elementary matrix we have, where R is rref,

$$E_s \ldots E_1 A = R. \tag{3}$$

By Theorem 1 each E_i has an inverse that is an elementary matrix. If we let $\tilde{E}_i$ denote its inverse and multiply equation (3) on both sides by $\tilde{E}_1 \ldots \tilde{E}_{s-1} \; \tilde{E}_s$ we obtain equation (1).

Proof of (b). If A has an inverse, then R is the identity in equation (1) and so we obtain equation (2). End of proof.

From the example before Theorem 2, we have that $(E_1)^{-1} \ldots (E_8)^{-1} = A$.

Elementary Matrices and Determinants

The determinants of elementary matrices are very easy to compute and together with Theorem 2 can be used to prove some very important properties of nxn matrices and their determinants. The main theorem is

Theorem 3. The determinant of an elementary matrix E is given as follows:

(a). det (E) = -1, if E interchanges two rows.

(b). det (E) = c, if E multiplies a row by a non-zero constant c.

(c). det (E) = 1, if If E multiplies row i by c and adds it to row j.

Proof of (a). This follows from Theorem 3.2 in the text. However, it is easy to prove it based on the fact that $\det(I_n) = a_{11} a_{22} \ldots a_{nn}$, where $a_{ii} = 1$, so when we interchange rows i and j, the term $a_{11} a_{22} \cdots a_{ji} \cdots a_{ij} \cdots a_{nn}$ is the determinant of E except possibly for its sign. But we see that the columns have been permuted and thus the sign would change to -1.

Proof of (b). E is just the same as I_n except for a non-zero constant c in place of a_{ii} to give $\det(E) = a_{11} a_{22} \ldots c \ldots aa_{nn} = c$ since all of the other terms equal 1.

Proof of (c). The elementary matrix E is that same as I_n except for one element c that is not on the main diagonal. This value c is multiplied by a 0 in any term that appear in the determinant of E since the only other non-zero terms in E are the 1's on the main diagonal. End of theorem.

The proof above will be clearer if you take a few minutes to compute the determinants of typical 3x3 elementary matrices.

Theorem 4. If E is an elementary matrix and B is an nxn matrix,

then $\det(E\,B) = \det(E)\det(B)$.

Proof. Case 1. If E interchanges two rows, then by Theorem 3.2 in the text, $\det(E\,B) = -\det(B)$.

Since det(E) = -1 by Theorem 3, then det(E B) = det(E) det(B).

Case 2. If E multiplies a row by a non-zero constant, then by Theorem 3.5 in the text, det(E B) = c det(B). Again, by Theorem 3, det(E) = c, so det(EB) = det(E) det(B).

Case 3. If E multiplies row i by c and adds it to row j, then by Theorem 3.6 in the text, det(E B) = det(B). Again, by Theorem 3, det(E) = 1, so det(E B) = det(E) det(B).

Elementary matrices can now be used to prove that det(AB) = det(A) det(B).

Theorem 5. (Theorem 3.8 in the text).

If A and B are nxn matrices, then det(AB) = det(A) det(B).

Proof. We will give part of the proof in a special case where A and B are non-singular and leave the remainder of the proof to the reader. Suppose A and B are non-singular. Then by Theorem 2(b) $A = \tilde{E}_1 \ldots \tilde{E}_s$ and $B = \hat{E}_1 \ldots \hat{E}_t$ and so

$$AB = \tilde{E}_1 \ldots \tilde{E}_s \hat{E}_1 \ldots \hat{E}_t$$

where all of the E's are elementary matrices. By Theorem 4,

$$\det(AB) = \det(\tilde{E}_1) \det(\tilde{E}_2 \ldots \tilde{E}_s \hat{E}_1 \ldots \hat{E}_t)$$

We can repeat using Theorem 4 to obtain

$$\det(AB) = \det(\tilde{E}_1) \det(\tilde{E}_2) \ldots \det(\tilde{E}_s) \det(\hat{E}_1) \ldots \det(\hat{E}_t)$$

But then it is easy to show that $\det(A) = \det(\tilde{E}_1) \det(\tilde{E}_2) \ldots \det(\tilde{E}_s)$ and

$\det(B) = \det(\hat{E}_1) \ldots \det(\hat{E}_t)$ which proves the result when A and B are non-singular.

Exercises

1. a. Determine the elementary matrix for each step in Example 5 in Sec. 1.6 of the text.

b. Show A^{-1} as a product of elementary matrices and verify by multiplying out the matrices that the product is the inverse of A.

c. Show A as a product of elementary matrices and verify by multiplying out the matrices that the product is A.

2. a. Determine the elementary matrix for each step in Example 6 in Sec. 1.6 of the text.

b. Verify Theorem 2(a) for A.

3. Show that Theorem 3 holds for the elementary matrices $E_1 \ldots E_8$ in the main example in this handout.

4. Compute the determinant of A in the main example in this handout by using only elementary matrices in conjunction with the Theorems in this handout.