Elementary Matrices

Ryan C. Daileda

Trinity University

Linear Algebra

Introduction

We defined matrix multiplication by the rule

$$AB = A\begin{pmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p\end{pmatrix} = \begin{pmatrix}A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p\end{pmatrix}.$$

That is, to compute AB we left multiply the columns of B by A. We will call this the *left multiplication rule*.

It turns out that we can also compute AB by instead right multiplying the *rows* of A by B.

This *right multiplication rule* will enable us to implement row operations as (left) multiplication by appropriate *elementary matrices*.

The Row-Matrix Product

Let A be an $m \times n$ matrix and let $\mathbf{v} \in \mathbb{R}^m$. Then $A^T\mathbf{v} \in \mathbb{R}^n$.
Let R_i denote the ith row of A (which is a $1 \times n$ matrix). Then

$$A^T\mathbf{v} = \begin{pmatrix} R_1^T & R_2^T & \cdots & R_m^T \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} = \sum_i v_i R_i^T = \left(\sum_i v_i R_i \right)^T$$

Taking the transpose of both sides we obtain

$$\sum_i v_i R_i = (A^T\mathbf{v})^T = \mathbf{v}^T A = \begin{pmatrix} v_1 & v_2 & \cdots & v_m \end{pmatrix} \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix}.$$

That is, $\mathbf{v}^T A$ is the linear combination of the *rows* of A, using the entries of $\mathbf{v}$ as weights.

Notice that we have

$$\begin{aligned}
AB &= (B^T A^T)^T = \left(B^T \begin{pmatrix} R_1^T & R_2^T & \cdots & R_m^T \end{pmatrix}\right)^T \\
&= \begin{pmatrix} B^T R_1^T & B^T R_2^T & \cdots & B^T R_m^T \end{pmatrix}^T \\
&= \begin{pmatrix} (R_1 B)^T & (R_2 B)^T & \cdots (R_m B)^T \end{pmatrix}^T = \begin{pmatrix} R_1 B \\ R_2 B \\ \vdots \\ R_m B \end{pmatrix}.
\end{aligned}$$

This shows that AB can be computed by letting B act on the *rows* of A (from the right).

The Right Multiplication Rule

This proves:

Theorem 1 (The Right Multiplication Rule)

Let A and B be (compatible) matrices, and let R_i denote the ith row of A. Then the ith row of AB is R_iB, which is the linear combination of the rows *of B using the entries of the ith row of A as weights.*

Remarks.

- This is similar to, though not quite identical to, what the book calls the Column-Row Expansion of AB.
- The name is meant to indicate that in the product AB we are thinking of B acting on the rows of A on the right.

Example

Suppose

$$A = \begin{pmatrix} 2 & -1 & 3 & 1 \\ 0 & 4 & 1 & -1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 2 & 2 \\ 0 & 3 & -1 \\ -1 & 2 & 3 \end{pmatrix}.$$

The right multiplication rule then gives

$$AB = \begin{pmatrix} 2\begin{pmatrix}1 & 0 & -1\end{pmatrix} - \begin{pmatrix}2 & 2 & 2\end{pmatrix} + 3\begin{pmatrix}0 & 3 & -1\end{pmatrix} + \begin{pmatrix}-1 & 2 & 3\end{pmatrix} \\ 0\begin{pmatrix}1 & 0 & -1\end{pmatrix} + 4\begin{pmatrix}2 & 2 & 2\end{pmatrix} + \begin{pmatrix}0 & 3 & -1\end{pmatrix} - \begin{pmatrix}-1 & 2 & 3\end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 9 & -4 \\ 9 & 9 & 4 \end{pmatrix},$$

which agrees with the usual “left” computation of AB (check it!).

Elementary Matrices

We can use the right-multiplication law to cast a new light on elementary row operations.

Let A be an $m \times n$ matrix. Since $\mathbf{e}_j$ is the jth columns of I, $\mathbf{e}_i^T$ is the ith row of $I^T = I$.

Because $\mathbf{e}_i$ has a 1 in the ith position and zeros elsewhere,

$$\mathbf{e}_i^T A = i\text{th row of } A.$$

Let E be the $n \times n$ matrix obtained by interchanging the ith and jth row of I.

The right-multiplication law then tells us that EA is the matrix obtained by interchanging the ith and jth row of A.

That is:

If E is the matrix obtained by interchanging the ith and jth rows of I, then interchanging the ith and jth rows of A can be achieved algebraically by left multiplication by E.

Example. Left multiplication by

$$E = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

will interchange the first and third rows of any $4 \times n$ matrix.

Now suppose E' is the matrix obtained from I by adding a times its kth row to its jth row. Then

$$i\text{th row of } E' = \begin{cases} \mathbf{e}_i^T & \text{if } i \neq j, \\ a\mathbf{e}_k^T + \mathbf{e}_j^T & \text{if } i = j. \end{cases}$$

Thus

$$i\text{th row of } E'A = \begin{cases} \mathbf{e}_i^T A & \text{if } i \neq j, \\ a(\mathbf{e}_k^T A) + \mathbf{e}_j^T A & \text{if } i = j. \end{cases}$$

$$= \begin{cases} i\text{th row of } A & \text{if } i \neq j, \\ a(k\text{th row of } A) + (j\text{th row of } A) & \text{if } i = j. \end{cases}$$

Thus:

> If E' is the matrix obtained from I by adding a times its kth row to its jth rows, then the same row operation can be performed on A by left multiplication by E'.

Example. If A is a $4 \times n$ matrix, the row operation

$$3R_4 + R_2 \to R_2$$

is given by left multiplication by

$$E' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

An entirely similar computation shows that:

If E'' is the matrix obtained from I by scaling its ith row by $a \neq 0$, then the same row operation can be performed on A by left multiplication by E''.

Example. Left multiplication by

$$E'' = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & -7 & \\ & & & & 1 \end{pmatrix}$$

will scale the fourth row of any $5 \times n$ matrix by -7.

Elementary Matrices

Matrices of the form E, E' or E'' (which are obtained from I by performing just a single row operation) are called *elementary matrices*.

Our work above proves:

Theorem 2

If A has the echelon form U, then there is a sequence of elementary matrices $E_1, E_2, \ldots, E_r$ (corresponding to the row operations used to transform A into U) so that

$$U = E_r E_{r-1} \cdots E_1 A.$$

Inverses of Elementary Matrices

If E is an elementary matrix corresponding to a certain elementary row operation, and E' is the matrix corresponding to the inverse operation, then

$$E'E = E'EI = E'(EI) = I,$$

since E' undoes what E does.

This proves:

Lemma 1

If E is an elementary matrix corresponding to a certain row operation, then E is invertible and E^{-1} is the matrix corresponding to the inverse operation.

Elementary Matrices and Invertibility

We can now use Theorem 2 to deduce another characterization of invertible matrices.

Corollary 1

A square matrix is invertible if and only if it is equal to a product of elementary matrices.

Proof. Because the product of invertible matrices is invertible, Lemma 1 implies that any product of elementary matrices is invertible.

For the converse, suppose A is invertible. Then A has a pivot in every row and column, so that

$$A \xrightarrow{\text{RREF}} I.$$

By Theorem 2, there are elementary matrices $E_1, E_2, \ldots, E_r$ so that

$$E_r E_{r-1} \cdots E_1 A = I.$$

Thus

$$A = (E_r E_{r-1} \cdots E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_r^{-1},$$

which by Lemma 1 is a product of elementary matrices. □