06. Elementary Matrices, my dear Watson

6.1 Definition

An *elementary matrix* of size $N \times N$ is a matrix obtained from the $N \times N$ Identity matrix $Id_{N\times N}$ by applying a single row operation.

For example, among the matrices on the adjoining page, all are elementary matrices, except for (iv). Matrix (i) is obtained by taking the 4×4 Identity matrix and switching rows 1 and 4; (ii) is obtained by rescaling row 3 of $Id_{3\times3}$ by the scalar 5; (iii) is obtained by taking $Id_{3\times3}$ and adding $\sqrt{2}$ times row 2 to row 3; (v) is obtained by taking $Id_{3\times3}$ and adding π times row 3 to row 1. (Note that we are doing linear algebra over the real numbers, $\mathbb{R}$, and later over the complex numbers $\mathbb{C}$). Hence, even though many of our examples involve only whole numbers (integers, $\mathbb{Z}$), real numbers are fair game.

Why is the matrix in (iv) *not* an elementary matrix? That matrix may be obtained from $Id_{3\times3}$ by scaling the second row by zero, but scaling by zero is not a *permitted* row operation, as we discussed previously.

To elaborate, call the matrix in (iv) B. Then the homogeneous equation $(Id_{3\times3})\vec{x} = \vec{0}$ has only one solution, while the equation $B\vec{x} = \vec{0}$ has many solutions. Our portfolio of *permitted* row operations is designed to preserve solution sets. `[Reader: elaborate farther. Something should be said about how row operations affect right hand sides. Say it]`

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

(ii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$

(iii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \sqrt{2} & 1 \end{pmatrix}$

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

(v) $\begin{pmatrix} 1 & 0 & \pi \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

> *Elementary, my dear Watson.*
>
> A much quoted phrase, ostensibly from the Sherlock Holmes stories of Arthur Conan Doyle, intended to convey the ease and simplicity with which Sherlock could solve mysteries.
>
> The phrase never actually appeared in the stories, but has become a part of popular culture.

[Continued on next page.]

6.2 Let your matrix do the row reducing

Proposition. *Every row operation may be manifested by left multiplication by an elementary matrix.*

We'll parse this claim in detail. First, what is meant by *manifested*? A dictionary may help (or maybe not, but give it a chance). Suppose we have a matrix M and we'd like to perform a particular row operation to it. Then we can find an elementary matrix E so that the matrix product $E \cdot M$ is precisely the result of the row operation we had in mind.

For example, say M is the matrix

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix},$$

and we would like to perform the row operation which adds (-9) times row 1 to row 3, or $III \longrightarrow III - 9I$ in Roman numeral notation. Then we can employ the 3×3 matrix

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -9 & 0 & 1 \end{pmatrix}.$$

The reader may verify that

$$\begin{aligned} E \cdot M &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -9 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & -8 & -16 & -24 \end{pmatrix}, \end{aligned}$$

which is the result of our intended row operation. How did we come up with the matrix E? We performed the same row operation on the 3×3 Identity matrix $Id_{3\times3}$. (The reader may ponder why this works.)

Suppose K is the 4×2 matrix below and we wish to scale its second row by the factor of $(1/2)$:

$$K \equiv \begin{pmatrix} 1 & 2 \\ 6 & 8 \\ 5 & 6 \\ 7 & 8 \end{pmatrix}.$$

We can manifest the scaling by left-multiplying by the 4×4 matrix below, which shall remain nameless:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 6 & 8 \\ 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix}.$$

How did we come up with this nameless 4×4 elementary matrix? You guessed it: we scaled the second row of $Id_{4\times4}$ by $\frac{1}{2}$.
What if we changed our mind and, instead, wanted to switch rows 2 and 4 of matrix K? This may be manifested by another 4×4 elementary matrix, multiplying on the left:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 6 & 8 \\ 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 7 & 8 \\ 5 & 6 \\ 6 & 8 \end{pmatrix}.$$

Following the previous ideas, we obtained the elementary 4×4 matrix to manifest the swapping by performing the given swap on the rows of the matrix $Id_{4\times4}$.

Do these examples prove the claimed proposition? Not formally. In principle, there could be other matrices and row operations that cannot be so manifested. But in practice, all the ideas needed for the general case are already present here. The reader is invited to write a general, formal argument, illustrated with expressions using plenty of the symbols $\ldots, \vdots, \cdots$, and $\ddots$ (ellipses and their variants).

6.3 RREF by matrix product

Theorem. *Let M be a $p \times q$ matrix. Then there is a sequence of elementary matrices of size $p \times p$, $E_1, E_2, \ldots, E_k$, so that the row reduced echelon form of M, $RREF(M)$ may be presented as*

$$E_k \cdot E_{k-1} \ldots E_3 \cdot E_2 \cdot E_1 \cdot M = RREF(M).$$

Thus the $RREF(M)$ may be obtained from M by multiplying on the left by a series of elementary matrices.

Proof. We know that every matrix can be put into RREF by a series of row operations. Each row operation may be "manifested" by left multiplication by an elementary matrix, as observed above. So the series of row operations can be manifested by a series of left multiplications by elementary matrices. $\square$

One may ask: why bother with this product of matrices presentation of our row operations? Is the list of row operations performed not sufficient? That's a good question. Matrix notation is found throughout mathematics while row operations are uncommon outside the context in which we used them. Also, matrix representation and multiplication can lead to theoretic results which are of interest in advanced variants of the subject. The ideas that surfaced here will be useful in our later discussions as well.