1.8 Row operations and matrix multiplication.

0. *Assumed background.*

   - 1.2 *Matrix multiplication.*
   - 1.7 *Row operations on matrices.*

Abstract. We introduce:—

- the notion of row operations matrices,
- the 'equivalence' of application of row operations and left-multiplication by row operation matrices.

1. **Definition. ('Standard base' for a 'vector space of matrices'.)**

   Fix any positive integers p, q.

   For each $i = 1, \cdots, p$, $j = 1, \cdots, q$, we define the $(p \times q)$-matrix $E_{i,j}^{p,q}$ to be the $(p \times q)$-matrix whose (i, j)-th entry is 1 and whose other entries are all 0.

   Remark. According to definition, there are altogether pq matrices $E_{i,j}^{p,q}$ as i, j vary. They are collectively referred to as the **'standard base' for the 'vector space of $(p \times q)$-matrices'**.

2. **Example (1). ('Standard base' for various 'vector spaces of matrices'.)**

   (a) $E_{1,1}^{2,3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $E_{1,2}^{2,3} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $E_{1,3}^{2,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$,

   $E_{2,1}^{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $E_{2,2}^{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $E_{2,3}^{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

   (b) $E_{1,1}^{3,3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $E_{1,2}^{3,3} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $E_{1,3}^{3,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$,

   $E_{2,1}^{3,3} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $E_{2,2}^{3,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $E_{2,3}^{3,3}$ $\cancel{E_{2,3}^{2,3}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$,

   $E_{3,1}^{3,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $E_{3,2}^{3,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $E_{3,3}^{3,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

3. **Lemma (1).**

   Let p, q be positive integers. Suppose s, t are integers between 1 and p.

   Let A be a $(p \times q)$-matrix, whose (i, j)-th entry is denoted by a_{ij}.

   Then $E_{s,t}^{p,p} A$ is the $(p \times q)$-matrix whose s-th row is $\begin{bmatrix} a_{t1} & a_{t2} & \cdots & a_{tq} \end{bmatrix}$, and whose every other entry is 0.

   Remark. In plain words, multiplying $E_{s,t}^{p,p}$ to A from the left results in a new $(p \times q)$-matrix in which:—

   - the s-th row is formed by the t-th row of A, and
   - every other row is 'set' to zero.

4. **Example (2). (Illustrations of Lemma (1).)**

   (a) Suppose A is the (3×4)-matrix whose (i, j)-th entry is given by a_{ij}. Then:

   i. $E_{1,2}^{3,3} A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

   ii. $E_{3,1}^{3,3} A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix}$.

   $E_{3,3}^{3,3}$

   iii. $\cancel{E_{3,1}^{3,3}} A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$

   (b) Suppose A is the (4×6)-matrix whose (i, j)-th entry is given by a_{ij}. Then:

   i. $E_{2,4}^{4,4} A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

ii. $E_{3,2}^{4,4}A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$

iii. $E_{4,1}^{4,4}A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \end{bmatrix}.$

5. **Proof of Lemma (1).**

For convenience, denote the (g, h)-th entry of $E_{s,t}^{p,p}$ by ε_{gh}.

So $\varepsilon_{gh} = \begin{cases} 1 & \text{if} \quad g = s \text{ and } h = t \\ 0 & \text{if} \quad g \neq s \text{ or } h \neq t \end{cases}$

By definition, the (g, k)-th entry of $E_{s,t}^{p,p}A$ is given by

$$\varepsilon_{g1}a_{1k} + \varepsilon_{g2}a_{2k} + \cdots + \varepsilon_{st}a_{tk} + \cdots + \varepsilon_{gp}a_{pk}.$$

- [We first focus on the s-th row of $E_{s,t}^{p,p}A$.]

  For each $k = 1, 2, \cdots, q$, the (s, k)-th entry of $E_{s,t}^{p,p}A$ is the product of the s-th row of $E_{s,t}^{p,p}$ and the k-th column of A, and therefore is given by

  $$\underbrace{\varepsilon_{s1}a_{1k} + \varepsilon_{s2}a_{2k} + \cdots + \varepsilon_{st}a_{tk} + \cdots + \varepsilon_{sp}a_{pk}}_{\text{all terms being 0 except possibly the term involving } \varepsilon_{st}} = \varepsilon_{st}a_{tk} = a_{tk}.$$

  Hence the s-th row of $E_{s,t}^{p,p}A$ is $[\ a_{t1} \quad a_{t2} \quad \cdots \quad a_{tq}\ ]$.

- [We now turn to every other row of $E_{s,t}^{p,p}A$.]

  Whenever $g \neq s$, we have $\varepsilon_{gh} = 0$ for each h. Then, no matter which k is, the (g, k)-th entry of $E_{s,t}^{p,p}A$ is a sum of p copies of 0's, and hence is 0.

6. **Lemma (2).**

Let A be a matrix with p rows. Let i, k be integers between 1 and p.

(a) *For any number α, the application of the row operation $\alpha R_i + R_k$ on A results in $(I_p + \alpha E_{k,i}^{p,p})A$.*

(b) *For any non-zero number β, the application of the row operation βR_k on A results in $(I_p + (\beta - 1)E_{k,k}^{p,p})A$.*

(c) *The application of the row operation $R_i \leftrightarrow R_k$ on A results in $(I_p - E_{i,i}^{p,p} - E_{k,k}^{p,p} + E_{i,k}^{p,p} + E_{k,i}^{p,p})A$.*

Proof of Lemma (2). Exercise. (Straightforward calculation with the help of Lemma (1).)

7. **Example (3). (Illustrations of Lemma (2).)**

Suppose A is the (3×4)-matrix whose (i, j)-th entry is given by a_{ij}. Then:

(a)

$$A \xrightarrow{4R_2+R_1} \begin{bmatrix} 4a_{21} + a_{11} & 4a_{22} + a_{12} & 4a_{23} + a_{13} & 4a_{24} + a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

$$= 4\begin{bmatrix} a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = 4E_{1,2}^{3,3}A + A = (I_3 + 4E_{1,2}^{3,3})A$$

(b)

$$A \xrightarrow{5R_2} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 5a_{21} & 5a_{22} & 5a_{23} & 5a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} + 4\begin{bmatrix} 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & 0 \end{bmatrix} = A + 4E_{2,2}^{3,3}A = (I_3 + 4E_{2,2}^{3,3})A$$

(c)

$$
\begin{aligned}
A \xrightarrow{R_1 \leftrightarrow R_3} & \begin{bmatrix} a_{31} & a_{32} & a_{33} & a_{34} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix} \\
= & \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} - \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \\
& + \begin{bmatrix} a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix} \\
= & A - E_{1,1}^{3,3}A - E_{3,3}^{3,3}A + E_{1,1}^{3,3}A + E_{3,3}^{3,3}A = (I_3 - E_{1,1}^{3,3} - E_{3,3}^{3,3} + E_{1,3}^{3,3} + E_{3,3}^{3,3})A.
\end{aligned}
$$

8. **Symbols and labels associated with Lemma (2).**

In symbolic terms, what we have described in Lemma (2) is the validity, for every matrix with p rows, say, A, of:—

- $A \xrightarrow{\alpha R_i + R_k} (I_p + \alpha E_{k,i}^{p,p})A$.
- $A \xrightarrow{\beta R_k} (I_p + (\beta - 1)E_{k,k}^{p,p})A$.
- $A \xrightarrow{R_i \leftrightarrow R_k} (I_p - E_{i,i}^{p,p} - E_{k,k}^{p,p} + E_{i,k}^{p,p} + E_{k,i}^{p,p})A$.

$\Rightarrow$ Existence of · $M[\rho]$

Because of their behaviour, it makes sense to give special labels for these matrices. From now on:—

- We label $I_p + \alpha E_{k,i}^{p,p}$ as $M[\alpha R_i + R_k]$.
- We label $I_p + (\beta - 1)E_{k,k}^{p,p}$ as $M[\beta R_k]$.
- We label $I_p - E_{i,i}^{p,p} - E_{k,k}^{p,p} + E_{i,k}^{p,p} + E_{k,i}^{p,p}$ as $M[R_i \leftrightarrow R_k]$.

Despite their seemingly complex formulae, these matrices are easy to write down explicitly, because they are the resultants of applications of the respective row operations on the identity matrix:—

- $I_p \xrightarrow{\alpha R_i + R_k} I_p + \alpha E_{k,i}^{p,p} = M[\alpha R_i + R_k]$.
- $I_p \xrightarrow{\beta R_k} I_p + (\beta - 1)E_{k,k}^{p,p} = M[\beta R_k]$.
- $I_p \xrightarrow{R_i \leftrightarrow R_k} I_p - E_{i,i}^{p,p} - E_{k,k}^{p,p} + E_{i,k}^{p,p} + E_{k,i}^{p,p} = M[R_i \leftrightarrow R_k]$.

$A \xrightarrow{\alpha R_i + R_k} M[\alpha R_i + R_k] \cdot A$

$I_p \xrightarrow{\alpha R_i + R_k} M[\alpha R_i + R_k] \cdot I_p$
$= M[\alpha R_i + R_k]$

9. **Definition. (Row-operation matrices.)**

   (a) *Let ρ be a row operation on matrices with p rows. The matrix $M[\rho]$ is called the* **row-operation matrix** *(or* **elementary matrix***)* **associated with** *ρ.*

   (b) *A* **row-operation matrix** *(or* **elementary matrix***)* **of size** *p is a $(p \times p)$-square matrix which is the row-operation matrix associated with some row-operation ρ on matrices with p rows.*

10. **Example (4). (Illustrations on row operation matrices.)**

Question: ① Can you define column operations on matrices?
② ~~At~~ Can you associate matrices to them?

   (a) For the row operation $3R_4 + R_2$ on matrices with 5 rows, its row operation matrix $M[3R_4 + R_2]$ is given by

$$I_5 \xrightarrow{3R_4 + R_2} M[3R_4 + R_2] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

   (b) For the row operation $4R_3$ on matrices with 5 rows, its row operation matrix $M[4R_3]$ is given by

$$I_5 \xrightarrow{4R_3} M[R_4 \leftrightarrow R_3] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

☆ Question: Why do we only focus on row operations now?

(c) For the row operation $R_2 \leftrightarrow R_5$ on matrices with 5 rows, its row operation matrix $M[R_2 \leftrightarrow R_5]$ is given by

$$I_5 \xrightarrow{R_2 \leftrightarrow R_5} M[R_2 \leftrightarrow R_5] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

11. **Theorem (3). ('Dictionary' between row operations and matrix multiplication from the left.)**

Let p, q be positive integers.

For any row operation ρ on matrices with p rows, there exists some unique $(p \times p)$-square matrix H, namely, $H = M[\rho]$, such that for any matrix with p rows, say, A, the matrix HA is the resultant of the application of ρ on the matrix A.

Remark. Theorem (3) below describes a 'dictionary' between the collection of all row operations on matrices with p rows and the collection of all row-operation matrices of size p.

This 'dictionary' tells us the 'application of row operations' and the 'multiplication from the left by row-operation matrices' are two ways of thinking about the same thing.

To write

$$\text{'the row operation } A \xrightarrow{\rho} A' \text{ is valid'}$$

$\Rightarrow$ $A \xrightarrow{\rho} HA$
$\qquad\qquad\qquad \| $
$\qquad\qquad\quad M[\rho]A$

is the same as writing

$$\text{'the equality } A' = M[\rho]A \text{ holds'.}$$

12. **Proof of Theorem (3).**

The 'existence' part of the argument follows immediately from Lemma (2). We only need to handle the 'uniqueness' part of the argument.

Suppose ρ is a row operation on $(p \times q)$ matrix, and H is a $(p \times p)$-square matrix for which the following holds:—

- For any $(p \times q)$-matrix A, the application of ρ on A results in HA.

Now, in particular, the application of ρ on I_p results in H.

Recall that by the definition of $M[\rho]$, the application of ρ on I_p results in $M[\rho]$.

It follows that $H = M[\rho]$.

$I_p \xrightarrow{\rho} H \cdot I_p = H$
$\qquad\quad \| \qquad\quad \|$
$\quad M[\rho]I_p = M[\rho]$

13. The 'dictionary' informs us that to apply a sequence of row operations on a matrix is the same as to multiply the product of the corresponding row-operation matrices (put in an appropriate order) to the same matrix from the left.

Theorem (4). (Corollary to Theorem (3).)

Let $A_1, A_2, \cdots, A_N$ be $(p \times q)$-matrices.

Suppose $A_1, A_2, \cdots, A_N$ are row-equivalent to each other, and A_1 is joint to A_N by some sequence of row operations $\rho_1, \rho_2, \cdots, \rho_{N-1}$:

$$A_1 \xrightarrow{\rho_1} A_2 \xrightarrow{\rho_2} A_3 \xrightarrow{\rho_3} \cdots \xrightarrow{\rho_{N-2}} A_{N-1} \xrightarrow{\rho_{N-1}} A_N$$

$M[\rho_1]A_1 \quad M[\rho_2]M[\rho_1]A_1 \quad \cdots$

Then the equalities

$$A_2 = M[\rho_1]A_1, \quad A_3 = M[\rho_2]A_2, \quad \cdots\cdots, \quad A_{N-1} = M[\rho_{N-2}]A_{N-2}, \quad A_N = M[\rho_{N-1}]A_{N-1},$$
$$\boxed{A_N = M[\rho_{N-1}]M[\rho_{N-2}] \cdot \ldots \cdot M[\rho_3]M[\rho_2]M[\rho_1]A_1}$$

$\Leftarrow$ be careful with the order of multiplications.

hold.

Proof of Theorem (4). This is an immediate consequence of Theorem (3).

14. **Comment on the content of Theorem (4).**

Implicit in this result is the information on:—

- how we can obtain the product
$$M[\rho_{N-1}]M[\rho_{N-2}] \cdots M[\rho_3]M[\rho_2]M[\rho_1]$$
without having to perform matrix multiplication.

This product is the resultant of the application of the sequence of row operations $\rho_1, \rho_2, \rho_3, \cdots, \rho_{N-2}, \rho_{N-1}$ on the identity matrix I_p:

$$\begin{aligned} I_p \quad & \xrightarrow{\rho_1} M[\rho_1] \xrightarrow{\rho_2} M[\rho_2]M[\rho_1] \xrightarrow{\rho_3} M[\rho_3]M[\rho_2]M[\rho_1] \\ & \xrightarrow{\rho_4} \cdots\cdots\cdots \xrightarrow{\rho_{N-2}} M[\rho_{_{N-2}}]\cdots M[\rho_3]M[\rho_2]M[\rho_1] \xrightarrow{\rho_{N-1}} M[\rho_{_{N-1}}]M[\rho_{_{N-2}}]\cdots M[\rho_3]M[\rho_2]M[\rho_1] \end{aligned}$$

15. **Example (5). (Illustrations on Theorem (4).)**

(a) The sequence of row operations below joins A and A'':

$$A = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 0 & 2 \end{bmatrix} \xrightarrow{1R_1+R_2} A' = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 0 & 0 & 2 \end{bmatrix} \xrightarrow{2R_2+R_1} A'' = \begin{bmatrix} 3 & 4 & 5 & 5 \\ 1 & 2 & 2 & 2 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

Then $A'' = H_2H_1A$, in which

$$H_1 = M[1R_1 + R_2] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad H_2 = M[2R_2 + R_1] = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

So $A'' = HA$, in which $H = H_2H_1 = \begin{bmatrix} 3 & 2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(b) The sequence of row operations below joins B and B'':

$$B = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 2 & -2 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix} \xrightarrow{4R_2} B' = \begin{bmatrix} 1 & 2 & 2 & -1 \\ 8 & -8 & 4 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix} \xrightarrow{-2R_1} B'' = \begin{bmatrix} -2 & -4 & -4 & 2 \\ 8 & -8 & 4 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

Then $B'' = H_2H_1B$, in which

$$H_1 = M[4R_2] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad H_2 = M[-2R_1] = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

So $B'' = HB$, in which $H = H_2H_1 = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(c) The sequence of row operations below joins C and C'':

$$C = \begin{bmatrix} 1 & 2 & 2 & 0 \\ 3 & 0 & 3 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} C' = \begin{bmatrix} 3 & 0 & 3 & 1 \\ 1 & 2 & 2 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} C'' = \begin{bmatrix} 3 & 0 & 3 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & 2 & 2 & 0 \end{bmatrix}.$$

Then $C'' = H_2H_1C$, in which $H_1 = M[R_1 \leftrightarrow R_2] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $H_2 = M[R_2 \leftrightarrow R_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

So $C'' = HC$, in which $H = H_2H_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

(d) The sequence of row operations below joins A_1 and A_4:

$$A_1 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 0 & 2 \end{bmatrix} \xrightarrow{1R_1+R_2} A_2 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 0 & 0 & 2 \end{bmatrix} \xrightarrow{2R_3} A_3 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 2 & 0 & 0 & 4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} A_4 = \begin{bmatrix} 2 & 0 & 0 & 4 \\ 1 & 2 & 2 & 2 \\ 1 & 0 & 1 & 1 \end{bmatrix}.$$

Then $A_4 = H_3H_2H_1A_1$, in which

$$H_1 = M[1R_1 + R_2] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad H_2 = M[2R_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad H_3 = M[R_1 \leftrightarrow R_3] = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

So $A_4 = HA_1$, in which $H = H_3H_2H_1 = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

16. **Theorem (5). (Matrix of 'reverse row-operation'.)**

Suppose ρ is a row operation on matrices with p rows, and $\widetilde{\rho}$ is its corresponding 'reverse row operation' on matrices with p rows.

Then $M[\widetilde{\rho}]M[\rho] = I_p$ and $M[\rho]M[\widetilde{\rho}] = I_p$.

$A \xrightarrow{\rho} A' \xrightarrow{\widetilde{\rho}} A$

$A' = M[\rho]A$ $\qquad M[\widetilde{\rho}]M[\rho]A$

Proof of Theorem (5).

Suppose ρ is a row operation on matrices with p rows, and $\widetilde{\rho}$ is its corresponding 'reverse row operation' on matrices with p rows.

By the definition of row-operation matrices, we have the sequence of row operations:

$\Rightarrow M[\widetilde{\rho}]M[\rho]A = A$ holds for any matrices with p rows

$$I_p \xrightarrow{\rho} M[\rho] \xrightarrow{\widetilde{\rho}} M[\widetilde{\rho}]M[\rho].$$

As $\widetilde{\rho}$ is the 'reverse row operation' for ρ, $M[\widetilde{\rho}]M[\rho] = I_p$.

Note that ρ is the 'reverse row operation' for $\widetilde{\rho}$. Repeating the argument, with the roles of $\rho, \widetilde{\rho}$ interchanged, we deduce that $M[\rho]M[\widetilde{\rho}] = I_p$.

$\Rightarrow M[\widetilde{\rho}]M[\rho] = I_p$.

Remark. The 'formulae' for the row-operation matrices of 'reverse row operations' for row operations of various types are explicitly described below:—

Row operation ρ on matrices with p rows.	Row-operation matrix $M[\rho]$.	'Reverse row operation' $\widetilde{\rho}$ for ρ.	Row-operation matrix $M[\widetilde{\rho}]$.
$\alpha R_i + R_k$.	$I_p + \alpha E^{p,p}_{k,i}$.	$-\alpha R_i + R_k$.	$I_p - \alpha E^{p,p}_{k,i}$.
βR_k.	$I_p + (\beta - 1)E^{p,p}_{k,k}$.	$(1/\beta)R_k$.	$I_p + (1/\beta - 1)E^{p,p}_{k,k}$.
$R_i \leftrightarrow R_k$.	$I_p - E^{p,p}_{i,i} - E^{p,p}_{k,k} + E^{p,p}_{i,k} + E^{p,p}_{k,i}$.	$R_i \leftrightarrow R_k$.	$I_p - E^{p,p}_{i,i} - E^{p,p}_{k,k} + E^{p,p}_{i,k} + E^{p,p}_{k,i}$.

17. **Example (6). (Row operation matrices and matrices of corresponding 'reverse row operations'.)**

(a) For the row operation $3R_4 + R_2$ on matrices with 5 rows, its row operation matrix $M[3R_4 + R_2]$ is given by

check that: $(I_p + \alpha E^{p,p}_{k,i})(I_p - \alpha E^{p,p}_{k,i}) = I_p$ ☆

$$I_5 \xrightarrow{3R_4+R_2} M[3R_4 + R_2] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The corresponding 'reverse row operation' on matrices with 5 rows is $-3R_4 + R_2$, and its row operation matrix $M[-3R_4 + R_2]$ is given by

$$I_5 \xrightarrow{-3R_4+R_2} M[-3R_4 + R_2] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -3 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(b) For the row operation $4R_3$ on matrices with 5 rows, its row operation matrix $M[4R_3]$ is given by

$$I_5 \xrightarrow{4R_3} M[R_4 \leftrightarrow R_3] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The corresponding 'reverse row operation' on matrices with 5 rows is $\frac{1}{4}R_3$, and its row operation matrix $M[\frac{1}{4}R_3]$ is given by

$$I_5 \xrightarrow{\frac{1}{4}R_3} M[\frac{1}{4}R_3] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(c) For the row operation $R_2 \leftrightarrow R_5$ on matrices with 5 rows, its row operation matrix $M[R_2 \leftrightarrow R_5]$ is given by

$$I_5 \xrightarrow{R_2 \leftrightarrow R_5} M[R_2 \leftrightarrow R_5] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

The corresponding 'reverse row operation' on matrices with 5 rows is $R_2 \leftrightarrow R_5$ itself, and its row operation matrix is $M[R_2 \leftrightarrow R_5]$ itself.

18. Because of the ‘dictionary’ between row operations and row-operation matrices (Theorem (3)):—

- whatever can be done through the application of row operations can also be done through the multiplication from the left by their corresponding row-operation matrices, and
- whatever can be done through the multiplication from the left by row-operation matrices can also be done through the application of the corresponding row operations which determine these row-operation matrices.

For this reason, you may wonder why we need to learn both.

It will transpire that dependent on the situation, to think (and work) with one of them may be more advantageous.

Below is an illustration of a problem which is easy to handle in terms of row operation matrices (and equalities amongst matrices), but not so easy in terms of row operations.

Illustration.

Let $C = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 3 & 3 \\ 2 & 4 & 0 \\ 1 & 3 & 2 \end{bmatrix}$.

We verify that C is not row-equivalent to D:

Argue by contradiction.

Suppose it were true that C was row-equivalent to D.

Then there would be some sequence of row operations joining C to D:

$$C \xrightarrow{\rho_1} \xrightarrow{\rho_2} \cdots \xrightarrow{\rho_{k-1}} \xrightarrow{\rho_k} D.$$

Denote the row-operation matrix of ρ_j by $M[\rho_j]$ for each j.

Then the equality

$$D = M[\rho_k]M[\rho_{k-1}] \cdot \ldots \cdot M[\rho_2]M[\rho_1]C$$

$D = HC.$

would hold.

Write $H = M[\rho_k]M[\rho_{k-1}] \cdot \ldots \cdot M[\rho_2]M[\rho_1]$.

$[d_1\ d_2\ d_3] = H[C_1\ C_2\ C_3]$

For each ℓ, denote the ℓ-th column of C, D by $\mathbf{c}_\ell, \mathbf{d}_\ell$ respectively. Then the equalities

$= [HC_1\ HC_2\ HC_3]$

$$\mathbf{d}_1 = H\mathbf{c}_1, \qquad \mathbf{d}_2 = H\mathbf{c}_2$$

would hold.

$\Rightarrow \begin{cases} d_1 = HC_1 \\ d_2 = HC_2 \end{cases}$

Note that $\mathbf{c}_2 = 2\mathbf{c}_1$. Then

$$\begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix} = \mathbf{d}_2 = H\mathbf{c}_2 = H(2\mathbf{c}_1) = 2H\mathbf{c}_1 = 2\mathbf{d}_1 = 2\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix},$$

which is impossible.

Hence, in the first place, C is not row-equivalent to D.

Thm (6) & Thm (7) are corollaries of previous theorems.

19. **Theorem (6). (Re-formulation of row-equivalence between matrices in terms of equalities for matrices.)**

Suppose A, B are matrices with p rows. Then the statements below are logically equivalent:—

(1) *A, B are row-equivalent to each other.*

(2) *There are (finitely many) row-operation matrices $G_1, G_2, \cdots, G_{k-1}, G_k$ such that $B = G_kG_{k-1} \cdot \ldots \cdot G_2G_1A$.*

Now suppose any one of the above holds (so that both hold). Then there are some row-operation matrices $H_1, H_2, \cdots, H_{k-1}, H_k$ such that:—

- *$A = H_1H_2 \cdot \ldots \cdot H_{k-1}H_kB$, and*
- *for each $j = 1, 2, \cdots, k$, the equalities $H_jG_j = I_p$ and $G_jH_j = I_p$.*

Proof of Theorem (6). Exercise. (Apply Theorem (3) and Theorem (5).)

Remark. This result is a useful device for theoretical discussions in the future, because we can make use of it to introduce equalities (instead of just row operations) in such discussions.

20. A seemingly trivial consequence of Theorem (6) is that:—

- *if a square matrix is row-equivalent to the identity matrix then the square matrix concerned is a product of row-operation matrices.*

In fact, much more can be said when we combine what is (implicitly) stated in Theorem (4) about products of row-operation matrices. This will become relevant when we introduce the notion of *invertibility* of square matrices.

Theorem (7).

Suppose A is a $(p \times p)$-square matrix. Then the statements below are logically equivalent:—

(1) *A is row-equivalent to I_p.*

(2) *A is a product of row-operation matrices.*

Proof of Theorem (7).

Suppose A is a $(p \times p)$-square matrix.

(a) By Theorem (6), if A is row-equivalent to I_p then A is a product of row-operation matrices.

(b) Suppose that A is a product of row-operation matrices, say, $H_1, H_2, \cdots, H_{k-1}, H_k$, and that the equality $A = H_1 H_2 \cdots H_{k-1} H_k$ holds.

Denote by $\gamma_1, \gamma_2, \cdots, \gamma_{k-1}, \gamma_k$ the respective row operations on $(p \times p)$-matrices to which the row-operation matrices $H_1, H_2, \cdots, H_{k-1}, H_k$ are associated.

We have $A = H_1 H_2 \cdots H_{k-1} H_k = H_1 H_2 \cdots H_{k-1} H_k I_p$.

Then we have the sequence of row operations

$$I_p \xrightarrow{\gamma_k} H_k \xrightarrow{\gamma_{k-1}} H_{k-1} H_k \xrightarrow{\gamma_{k-2}} \cdots\cdots\cdots \xrightarrow{\gamma_3} H_3 \cdots H_{k-1} H_k \xrightarrow{\gamma_2} H_2 H_3 \cdots H_{k-1} H_k \xrightarrow{\gamma_1} H_1 H_2 H_3 \cdots H_{k-1} H_k = A$$

Therefore I_p is row-equivalent to A. Hence A is row-equivalent to I_p.