3.4 Row-equivalence from the point of view of invertible matrices.

0. *Assumed background.*
   - Whatever has been covered in Topics 1-2, especially:—
     - ∗ 1.5 *Linear combinations.*
     - ∗ 1.6 *Linear dependence and linear independence.*
     - ∗ 1.8 *Row operations and matrix multiplication.*
     - ∗ 2.2 *Row-echelon forms and reduced row-echelon forms.*
   - 3.2 *Invertibility and row operations.*

Invertible matrices preserve linear combinations / linear dependence / linear independence

Abstract. We introduce:—

- how to visualize row-equivalence through left-multiplication by invertible matrices.
- how linear relations amongst various columns in a given matrix are 'preserved' upon application of row operations on the given matrix.
- the uniqueness of reduced row-echelon form which is row-equivalent to a given matrix.
- the notion of rank for an arbitrary matrix.

The proof of the uniqueness result on reduced row-echelon form row-equivalent to an arbitrarily given matrix is contained in the *appendix*, but the ideas in the proof is displayed in a 'concrete' example.

1. When we introduce the 'dictionary' between row operations and row-operation matrices, we discover the result below, labelled (⋆) here:

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

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

   $I_p + \alpha E^{p\times p}_{ij}$

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

   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:—

   Each G_i is invertible

   - *$A = H_1 H_2 \cdot \ldots \cdot H_{k-1} H_k B$, and*
   - *for each $j = 1, 2, \cdots, k$, the equalities $H_j G_j = I_p$ and $G_j H_j = I_p$.* $\Rightarrow H_j = G_j^{-1}$

   In terms of the re-formulation for the notion of invertibility in terms of row operations and row-operation matrices, we can re-formulate Theorem (⋆) as:—

2. **Theorem (1). (Re-formulation of row-equivalence in terms of multiplication by invertible matrices from the left.)**

   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 exists some invertible $(p \times p)$-square matrix G such that $B = GA$.*

   $G = G_k G_{k-1} \cdots G_2 G_1$

   Now suppose any one of the above holds (so that both hold). Then, for the same G, the equality $A = G^{-1}B$ holds.

   Remark. Such a re-formulation of row-equivalence is useful in theoretical discussions because through it, we can think and work in terms of equalities which are as 'simple' as possible.

3. **Example (1). (Illustrations on the content of Theorem (1).)**

   $A \xrightarrow{\rho_1} \xrightarrow{\rho_2} \xrightarrow{\rho_3} B$

   (a) Let $A = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 & 0 & 4 \\ 1 & 2 & 2 & 2 \\ 1 & 0 & 1 & 1 \end{bmatrix}$.

   $B = GA$

   A is row-equivalent to B under the sequence of row operations below:

   $G = M(\rho_3) M(\rho_2) M(\rho_1)$

   $$A \xrightarrow{1R_1+R_2} \xrightarrow{2R_3} \xrightarrow{R_1 \leftrightarrow R_3} B.$$

   $I_p \xrightarrow{\rho_1} \xrightarrow{\rho_2} \xrightarrow{\rho_3} G$

   Coincidentally, the equality $B = GA$ holds, in which G is the invertible (3×3)-square matrix given by

   $$I_3 \xrightarrow{1R_1+R_2} \xrightarrow{2R_3} \xrightarrow{R_1 \leftrightarrow R_3} G = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

(b) Let $A = \begin{bmatrix} 0 & -1 & -2 & 2 & 6 \\ 2 & -3 & 0 & 0 & 0 \\ -2 & 2 & -2 & 1 & 1 \\ 2 & -4 & -2 & 2 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 3 & 0 & 6 \\ 0 & 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

A is row-equivalent to B under the sequence of row operations below:

$$A \xrightarrow{R_1 \leftrightarrow R_4} \xrightarrow{\frac{1}{2}R_1} \xrightarrow{-2R_1+R_2} \xrightarrow{2R_1+R_3} \xrightarrow{1R_2+R_4} \xrightarrow{2R_2+R_3} \xrightarrow{-1R_3} \xrightarrow{2R_2+R_1} \xrightarrow{3R_3+R_1} \xrightarrow{2R_3+R_2} B.$$

Coincidentally, the equality $B = GA$ holds, in which G is the invertible (4×4)-square matrix given by

$$I_4 \xrightarrow{R_1 \leftrightarrow R_4} \xrightarrow{\frac{1}{2}R_1} \xrightarrow{-2R_1+R_2} \xrightarrow{2R_1+R_3} \xrightarrow{1R_2+R_4} \xrightarrow{2R_2+R_3} \xrightarrow{-1R_3} \xrightarrow{2R_2+R_1} \xrightarrow{3R_3+R_1} \xrightarrow{2R_3+R_2} G = \begin{bmatrix} 0 & -4 & -3 & 3/2 \\ 0 & -3 & -2 & 1 \\ 0 & -2 & -1 & 1 \\ 1 & 1 & 0 & -1 \end{bmatrix}.$$

(c) Let $A = \begin{bmatrix} 0 & 0 & 2 & 3 & 5 & -7 & 12 \\ -1 & 2 & 1 & -1 & 0 & -2 & 0 \\ 2 & -4 & -1 & 3 & 2 & 1 & 5 \\ 3 & -6 & -1 & 5 & 4 & 0 & 10 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -2 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

A is row-equivalent to B under the sequence of row operations below:

$$A \xrightarrow{R_1 \leftrightarrow R_2} \xrightarrow{-1R_1} \xrightarrow{-2R_1+R_3} \xrightarrow{-3R_1+R_4} \xrightarrow{R_2 \leftrightarrow R_3} \xrightarrow{-2R_2+R_3} \xrightarrow{-2R_2+R_4} \xrightarrow{1R_2+R_1} \xrightarrow{-2R_3+R_1} \xrightarrow{-1R_3+R_2} B.$$

Coincidentally, the equality $B = GA$ holds, in which G is the invertible (4×4)-square matrix given by

$$I_4 \xrightarrow{R_1 \leftrightarrow R_2} \xrightarrow{-1R_1} \xrightarrow{-2R_1+R_3} \xrightarrow{-3R_1+R_4} \xrightarrow{R_2 \leftrightarrow R_3} \xrightarrow{-2R_2+R_3} \xrightarrow{-2R_2+R_4}$$

$$\xrightarrow{1R_2+R_1} \xrightarrow{-2R_3+R_1} \xrightarrow{-1R_3+R_2} G = \begin{bmatrix} -2 & 9 & 5 & 0 \\ -1 & 6 & 3 & 0 \\ 1 & -4 & -2 & 0 \\ 0 & -1 & -2 & 1 \end{bmatrix}.$$

$rkA = rkB = 3$ rank = number of non-zero rows (for r-e-f r-r-e-f)

(d) Let $A = \begin{bmatrix} 2 & 7 & -8 \\ 1 & 4 & -5 \\ -1 & -1 & 1 \\ -2 & -6 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. $G = \underbrace{G_k G_{k-1} \cdots G_1}$

A is row-equivalent to B under the sequence of row operations below:

$$A \xrightarrow[G_1]{R_1 \leftrightarrow R_2} \xrightarrow{-2R_1+R_2} \xrightarrow{1R_1+R_3} \xrightarrow{2R_1+R_4} \xrightarrow{3R_2+R_3} \xrightarrow{2R_2+R_4} \xrightarrow{-1R_2} \xrightarrow{\frac{1}{2}R_3} \xrightarrow{-4R_2+R_1} \xrightarrow{-3R_3+R_1} \xrightarrow[G_k]{2R_3+R_2} B.$$

Coincidentally, the equality $B = GA$ holds, in which G is the invertible (4×4)-square matrix given by

same sequence of row operations

$$I_4 \xrightarrow{R_1 \leftrightarrow R_2} \xrightarrow{-2R_1+R_2} \xrightarrow{1R_1+R_3} \xrightarrow{2R_1+R_4} \xrightarrow{3R_2+R_3} \xrightarrow{2R_2+R_4}$$

$$\xrightarrow{-1R_2} \xrightarrow{\frac{1}{2}R_3} \xrightarrow{-4R_2+R_1} \xrightarrow{-3R_3+R_1} \xrightarrow{2R_3+R_2} G = \begin{bmatrix} -1/2 & 1/2 & -3/2 & 0 \\ 2 & -3 & 1 & 0 \\ 3/2 & -5/2 & 1/2 & 0 \\ 2 & -2 & 0 & 1 \end{bmatrix}.$$

$= G_k G_{k-1} \cdots G_1$

(e) Let $A = \begin{bmatrix} 1 & 1 & 1 & -2 & -3 & 1 \\ 2 & 2 & 2 & -7 & -8 & -3 \\ 3 & 2 & 1 & -5 & -7 & 5 \\ 2 & 4 & 6 & -4 & -9 & 2 \\ 0 & 1 & 2 & 0 & -2 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 4 \\ 0 & 1 & 2 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

A is row-equivalent to B under the sequence of row operations below:

$$A \xrightarrow{-2R_1+R_2} \xrightarrow{-3R_1+R_3} \xrightarrow{-2R_1+R_4} \xrightarrow{R_2 \leftrightarrow R_5} \xrightarrow{1R_2+R_3} \xrightarrow{-2R_2+R_4} \xrightarrow{3R_3+R_5} \xrightarrow{2R_4+R_5} \xrightarrow{-1R_2+R_1} \xrightarrow{2R_3+R_1} \xrightarrow{1R_4+R_1} \xrightarrow{2R_4+R_2} B.$$

Coincidentally, the equality $B = GA$ holds, in which G is the invertible (5×5)-square matrix given by

$$I_5 \xrightarrow{-2R_1+R_2} \xrightarrow{-3R_1+R_3} \xrightarrow{-2R_1+R_4} \xrightarrow{R_2 \leftrightarrow R_5} \xrightarrow{1R_2+R_3} \xrightarrow{-2R_2+R_4} \xrightarrow{3R_3+R_5} \xrightarrow{2R_4+R_5}$$

$$\xrightarrow{-1R_2+R_1} \xrightarrow{2R_3+R_1} \xrightarrow{1R_4+R_1} \xrightarrow{2R_4+R_2} G = \begin{bmatrix} -7 & 0 & 2 & 1 & -1 \\ -4 & 0 & 0 & 2 & -3 \\ -3 & 0 & 1 & 0 & 1 \\ -2 & 0 & 0 & 1 & -2 \\ -15 & 1 & 3 & 2 & -1 \end{bmatrix}.$$

4. We are going to introduce a few results about linear combinations, linear dependence and linear independence which can be applied in proving deep results. When seen from the point of view of invertible matrices, the results themselves look obvious. However, they will give rise to highly non-trivial results when they are re-interpreted in terms of row operations.

We start with a elementary result on 'matrix algebra':

Lemma (2). ('Preservation' of linear relations by left-multiplication by invertible matrices.)

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ be column vectors with p entries, and $\alpha_1, \alpha_2, \cdots, \alpha_n$ be numbers.

Suppose G is an invertible $(p \times p)$-square matrix.

Then $\boxed{\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n}$ if and only if $G\mathbf{v} = \alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n$.

Remark. As it will become apparent in the argument, the invertibility of G is needed in only one 'direction' of the conclusion, where we wish to 'cancel' G from both sides of an equality in which G is initially present. Here for the sake of simplicity, we state such a 'weaker' version of a more precise and general result.

5. **Proof of Lemma (2).**

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ be column vectors with p entries, and $\alpha_1, \alpha_2, \cdots, \alpha_n$ be numbers.

Suppose G is an invertible $(p \times p)$-square matrix.

(a) Suppose $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n$.
Then multiplying G to both sides of the above equality from the left, we obtain:

$$\begin{aligned} G\mathbf{v} &= G(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n) \\ &= \alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n \end{aligned}$$

distributive law

(b) Suppose $G\mathbf{v} = \alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n$.
By assumption, the matrix inverse G^{-1} of the matrix G is well-defined. Apply G^{-1}
Multiplying G^{-1} to both sides of the above equality from the left, we obtain:

$$\begin{aligned} \mathbf{v} = I_p\mathbf{v} = (G^{-1}G)\mathbf{v} = G^{-1}(G\mathbf{v}) &= G^{-1}(\alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n) \\ &= G^{-1}[G(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n)] \\ &= (G^{-1}G)(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n) \\ &= I_p(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n) \\ &= \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n \end{aligned}$$

6. Theorem (3), Theorem (4) and Theorem (5) are immediate consequences of Lemma (2).

Theorem (3). ('Preservation' of linear combinations by left-multiplication by invertible matrices.)

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ be column vectors with p entries.

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

(1) *$\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ with respect to scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$.*

(2) *$G\mathbf{v}$ is a linear combination of $G\mathbf{u}_1, G\mathbf{u}_2, \cdots, G\mathbf{u}_n$ with respect to scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$.*

Theorem (4). ('Preservation' of linear dependence and non-trivial linear relations by left-multiplication by invertible matrices.)

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be column vectors with p entries.

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

(1) *$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly dependent.* Find α_i ? not all zero $\alpha_1 u_1 + \cdots \alpha_n u_n = 0$

(2) *$G\mathbf{u}_1, G\mathbf{u}_2, \cdots, G\mathbf{u}_n$ are linearly dependent.*

Now suppose any one of the above holds (so that both hold).

Then, for any numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$ which are not all zero, the non-trivial linear relation $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}_p$ holds if and only if the non-trivial linear relation $\alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n = \mathbf{0}_p$ holds.

Theorem (5). ('Preservation' of linear independence by left-multiplication by invertible matrices.)

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be column vectors with p entries.

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

(1) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

(2) $G\mathbf{u}_1, G\mathbf{u}_2, \cdots, G\mathbf{u}_n$ *are linearly independent.*

Remark. We provide the proof for Theorem (5). That for Theorem (3) and Theorem (4) are left as exercises.

7. **Proof of Theorem (5).**

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be column vectors with p entries.

Suppose G is an invertible $(p \times p)$-square matrix.

(a) Suppose the statement (1) holds: $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

[We want to deduce that $G\mathbf{u}_1, G\mathbf{u}_2, \cdots, G\mathbf{u}_n$ *are linearly independent.*

This amounts to deducing: *For any numbers* $\alpha_1, \alpha_2, \cdots, \alpha_n$, *if* $\alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n = \mathbf{0}_p$ *then* $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.]

Pick any numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$. Suppose $\alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n = \mathbf{0}_p$.

Note that $\mathbf{0}_p = G\mathbf{0}_p$. Then $\alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n = G\mathbf{0}_p$.

Then, by Lemma (2), $\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_n \mathbf{u}_n = \mathbf{0}_p$.

Since $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are linearly independent, $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Hence the statement (2) holds.

(b) Suppose the statement (2) holds: $G\mathbf{u}_1, G\mathbf{u}_2, \cdots, G\mathbf{u}_n$ *are linearly independent.*

[We want to deduce that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent.*

This amounts to deducing: *For any numbers* $\alpha_1, \alpha_2, \cdots, \alpha_n$, *if* $\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_n \mathbf{u}_n = \mathbf{0}_p$ *then* $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.]

Pick any numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$. Suppose $\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_n \mathbf{u}_n = \mathbf{0}_p$.

Then by Lemma (2), we have $\alpha_1 G\mathbf{u}_1 + \alpha_2 G\mathbf{u}_2 + \cdots + \alpha_n G\mathbf{u}_n = G\mathbf{0}_p = \mathbf{0}_p$.

Since $G\mathbf{u}_1, G\mathbf{u}_2, \cdots, G\mathbf{u}_n$ are linearly independent, $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Hence the statement (1) holds.

8. **Re-interpretation of the results described by Theorem (3), Theorem (4), Theorem (5) in terms of row operations and row-equivalence.**

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ be column vectors with p entries.

Suppose $\mathbf{u}'_1, \mathbf{u}'_2, \cdots, \mathbf{u}'_n, \mathbf{v}'$ are row-equivalent to $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n, \mathbf{v}$ $\boxed{\text{under the same sequence of row operations.}}$

(a) According to Theorem (3):— $\quad [u_1 | u_2 | \cdots u_n | v] \xrightarrow{\text{row equiv}} [u_1' | u_2' | \cdots | u_n' | v']$

if $\mathbf{v}$ *is a linear combination of* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *with respect to scalars* $\alpha_1, \alpha_2, \cdots, \alpha_n$ *then* $\mathbf{v}'$ *is a linear combination of* $\mathbf{u}'_1, \mathbf{u}'_2, \cdots, \mathbf{u}'_n$ *with respect to scalars* $\alpha_1, \alpha_2, \cdots, \alpha_n$.

(b) According to Theorem (4):— $\quad [u_1 | u_2 | \cdots u_n | v] = G\,[u_1' | u_2' | \cdots | u_n' | v']$

if $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly dependent with the non-trivial relation* $\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_n \mathbf{u}_n = \mathbf{0}_p$, *then* $\mathbf{u}'_1, \mathbf{u}'_2, \cdots, \mathbf{u}'_n$ *are linear dependent with the non-trivial relation* $\alpha_1 \mathbf{u}'_1 + \alpha_2 \mathbf{u}'_2 + \cdots + \alpha_n \mathbf{u}'_n = \mathbf{0}_p$.

(c) According to Theorem (5):— $\quad = [G\cdot u_1' | G u_2' | \cdots | G\cdot u_n' | G\cdot v']$

if $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ *are linearly independent, then* $\mathbf{u}'_1, \mathbf{u}'_2, \cdots, \mathbf{u}'_n$ *are linear independent.*

$u_i = G\cdot u_i'$

Remark. These conclusions are not obvious (and not easy to prove) when we think in terms of row operations alone, because the 'equality symbol' does not appear in a relevant way in a discussion purely about row operations.

$v = G\cdot v'$

9. To give the intended application of Theorem (3), Theorem (4), Theorem (5) below, we embed the above re-interpretation into the context of relations amongst columns of row-equivalent matrices.

Theorem (6). (Corollary to Theorem (3), Theorem (4) and Theorem (5).)

Let B, C *be* $(p \times q)$*-matrix. Denote the* j*-th columns of* B, C *by* $\mathbf{b}_j, \mathbf{c}_j$ *respectively for each* $j = 1, 2, \cdots, q$.

Suppose B *is row-equivalent to* C. *Then the statements below hold:—*

(a) *If* $\mathbf{b}_j$ *is a linear combination of* $\mathbf{b}_{k_1}, \mathbf{b}_{k_2}, \cdots, \mathbf{b}_{k_n}$ *with respect to scalars* $\alpha_1, \alpha_2, \cdots, \alpha_n$ *then* $\mathbf{c}_j$ *is a linear combination of* $\mathbf{c}_{k_1}, \mathbf{c}_{k_2}, \cdots, \mathbf{c}_{k_n}$ *with respect to scalars* $\alpha_1, \alpha_2, \cdots, \alpha_n$.

(b) *If* $\mathbf{b}_{k_1}, \mathbf{b}_{k_2}, \cdots, \mathbf{b}_{k_n}$ *are linearly dependent with the non-trivial relation* $\alpha_1 \mathbf{b}_{k_1} + \alpha_2 \mathbf{b}_{k_2} + \cdots + \alpha_n \mathbf{b}_{k_n} = \mathbf{0}_p$, *then* $\mathbf{c}_{k_1}, \mathbf{c}_{k_2}, \cdots, \mathbf{c}_{k_n}$ *are linear dependent with the non-trivial relation* $\alpha_1 \mathbf{c}_{k_1} + \alpha_2 \mathbf{c}_{k_2} + \cdots + \alpha_n \mathbf{c}_{k_n} = \mathbf{0}_p$.

(c) *If* $\mathbf{b}_{k_1}, \mathbf{b}_{k_2}, \cdots, \mathbf{b}_{k_n}$ *are linearly independent, then* $\mathbf{c}_{k_1}, \mathbf{c}_{k_2}, \cdots, \mathbf{c}_{k_n}$ *are linear independent.*

10. Theorem (6) is the crucial piece of theoretical machinery needed in an argument for Theorem (7), about the uniqueness of reduced row-echelon form row-equivalent to an arbitrarily given matrix.

With Theorem (7) established, we can confirm the validity of Theorem (8), which has been stated earlier.

Theorem (7). (Uniqueness of reduced row-echelon form which is row-equivalent to a given matrix.)

Suppose that A is a matrix, and B, C are reduced row-echelon forms.

Further suppose that B is row-equivalent to A, and C is also row-equivalent to A.

Then $B = C$.

Theorem (8). (Existence and uniqueness of reduced row-echelon form which is row-equivalent to a given matrix.)

Suppose A is a matrix.

Then there exists some unique reduced row-echelon form A' such that A' is row-equivalent to A.

Remark. In the light of Theorem (8), it makes sense to write something like

'*the reduced row-echelon form which is row-equivalent to the matrix blah-blah-blah*'.

11. Now, in view of Theorem (8), it makes sense to introduce the notion of 'rank' for an arbitrary matrix.

Definition. (Rank of an arbitrary matrix.)

① If A, B are row equivalent, then rank of A = rank of B

Suppose A is a matrix. # of nonzero rows unique

Then the **rank** *of A is defined to be the rank of the reduced row-echelon form which is row-equivalent to A.*

Remark. This is one of several ways of formulating the definition for the notion of rank. They are, however, logically equivalent to this formulation.

② If G is invertible $A = G \cdot B$, then rank of A = the rank of B

12. The proof of Theorem (7) is contained in the *appendix*. The idea of the argument is illustrated in 'concrete terms' in Example (2) below.

In the argument for Theorem (7), we will need to use some observations about the 'linear relations' amongst the columns of reduced row-echelon forms. These observations are of interest on their own, and are stated below in the form of a theoretical result on its own. Example (2) will also serve as a 'concrete' illustration of the content of Theorem (9.)

Theorem (9). ('Linear relations' amongst columns of a reduced row-echelon form.)

Let C be a reduced row-echelon form with q columns. Denote the j-th column of C by $\mathbf{c}_j$ for each $j = 1, 2, \cdots, q$, and denote the (k, ℓ)-th entry of C by $\gamma_{k\ell}$ for each k, ℓ.

Denote the rank of C by r, and suppose the pivot columns of C, from left to right, are the d_1-th, d_2-th, ..., d_r-th columns of C.

Then:— (Pivot columns)

☆ $\mathrm{rk}\, C = r =$ number of pivot columns

(a) *$\mathbf{c}_{d_1}, \mathbf{c}_{d_2}, \cdots, \mathbf{c}_{d_r}$ are linearly independent.*

(b) *For each $j = 1, 2, \cdots, q$, if $\mathbf{c}_j$ is a free column, and the pivot columns strictly to the left of $\mathbf{c}_j$ are the d_1-th, d_2-th, ..., d_h-th columns, then $\mathbf{c}_j$ is a linear combination of $\mathbf{c}_{d_1}, \mathbf{c}_{d_2}, \cdots, \mathbf{c}_{d_h}$, with the linear relation $\mathbf{c}_j = \gamma_{1j}\mathbf{c}_{d_1} + \gamma_{2j}\mathbf{c}_{d_2} + \cdots + \gamma_{hj}\mathbf{c}_{d_h}$.*

(c) *For each $k = 1, 2, \cdots, r$, the d_k-th column of C, (which is the column vector $\mathbf{c}_{d_k}$,) is not a linear combination of the columns of C strictly to its left.*

In particular, the d_k-th column of C is not a linear combination of the d_1-th, d_2-th, ..., d_{k-1}-th columns of C.

Proof of Theorem (9). Exercise (on working with the definition of reduced row-echelon form).

13. **Example (2). (Illustration on the content of Theorem (9), and on the argument for Theorem (7).)**

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

☆ Number of free column

= number of nonzero rows

Denote the columns of C, from left to right, by $\mathbf{c}_j$ for each $j = 1, 2, \cdots, 9$.

definition of the rank.

(a) C is a reduced row-echelon form with rank 4.

The pivot columns of C are the 1-st, 3-rd, 4-th, 8-th columns, which are $\mathbf{c}_1, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_8$.

(b) We verify that These four column vectors

$$\mathbf{c}_1 = \begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix}, \quad \mathbf{c}_3 = \begin{bmatrix}0\\1\\0\\0\\0\\0\end{bmatrix}, \quad \mathbf{c}_4 = \begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}, \quad \mathbf{c}_8 = \begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}$$

are linearly independent:—

- Pick any numbers $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. Suppose $\alpha_1\mathbf{c}_1 + \alpha_2\mathbf{c}_3 + \alpha_3\mathbf{c}_4 + \alpha_4\mathbf{c}_8 = \mathbf{0}_6$.

  Then $\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\\\alpha_4\\0\\0\end{bmatrix} = \mathbf{0}_6$.

  Therefore $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$.

(c) The free columns of C are the 2-nd, 5-th, 6-th, 7-th, 9-th columns, which are the column vectors $\mathbf{c}_2, \mathbf{c}_5, \mathbf{c}_6, \mathbf{c}_7, \mathbf{c}_9$ respectively.

Note that:—

- $\mathbf{c}_2 = -2\mathbf{c}_1$,
- $\mathbf{c}_5 = 0 \cdot \mathbf{c}_1 + 1 \cdot \mathbf{c}_3 + 1 \cdot \mathbf{c}_4$,
- $\mathbf{c}_6 = 1 \cdot \mathbf{c}_1 - 2\mathbf{c}_3 - 1 \cdot \mathbf{c}_4$,
- $\mathbf{c}_7 = 1 \cdot \mathbf{c}_1 + 3\mathbf{c}_3 + 2\mathbf{c}_4$,
- $\cancel{\mathbf{c}_9 = 2\mathbf{c}_1 + 0 \cdot \mathbf{c}_3 + 3\mathbf{c}_4 + 0 \cdot \mathbf{c}_8,}$
- $\mathbf{c}_9 = -2\mathbf{c}_1 + 2\mathbf{c}_3 + 3\mathbf{c}_4 - 3\mathbf{c}_8$.

(d) Note that:—

- $\mathbf{c}_3$ is not a linear combination of $\mathbf{c}_1$.
- $\mathbf{c}_4$ is not a linear combination of $\mathbf{c}_1, \mathbf{c}_3$.
- $\mathbf{c}_8$ is not a linear combination of $\mathbf{c}_1, \mathbf{c}_3, \mathbf{c}_4$.

In fact, more can be said:—

- $\mathbf{c}_3$ is not a linear combination of $\mathbf{c}_1, \mathbf{c}_2$.
  (Reason: $\mathbf{c}_2$ is just a scalar multiple of $\mathbf{c}_1$.)
- $\mathbf{c}_4$ is not a linear combination of $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$.
  (Reason: $\mathbf{c}_2$ is just a linear combination of $\mathbf{c}_1, \mathbf{c}_3$.)
- $\mathbf{c}_8$ is not a linear combination of $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_5, \mathbf{c}_6, \mathbf{c}_7$.
  (Reason: Each of $\mathbf{c}_2, \mathbf{c}_5, \mathbf{c}_6, \mathbf{c}_7$ is just a linear combination of $\mathbf{c}_1, \mathbf{c}_3, \mathbf{c}_4$.)

(e) Let B be a (6×9)-matrix.

Denote the columns of B, from left to right, by $\mathbf{b}_j$ for each $j = 1, 2, \cdots, 9$.

Denote the $(k, \ell) - th$ entries of B by $\beta_{k\ell}$ for each k, ℓ.

Suppose B is a reduced row-echelon form, and suppose B is row-equivalent to C.

We want to verify that $B = C$:—

i. Recall that $\mathbf{c}_1, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_8$ are linearly independent.
   Then, since B is row-equivalent to C, $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_4, \mathbf{b}_8$ are linearly independent.

ii. In particular, $\mathbf{b}_1$ is linearly independent.
   Then $\mathbf{b}_1 \neq \mathbf{0}_6$.

   Since B is a reduced row-echelon form, $\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\\0\\0\\0\end{bmatrix} = \mathbf{c}_1$.

iii. Note that $\mathbf{c}_2 = -2\mathbf{c}_1$.
   Then, since B is row-equivalent to C, we have $\mathbf{b}_2 = -2\mathbf{b}_1 = -2\mathbf{c}_1 = \mathbf{c}_2$.

   So now $B = \begin{bmatrix} 1 & -2 & \beta_{13} & \beta_{14} & \beta_{15} & \beta_{16} & \beta_{17} & \beta_{18} & \beta_{19} \\ 0 & 0 & \beta_{23} & \beta_{24} & \beta_{25} & \beta_{26} & \beta_{27} & \beta_{28} & \beta_{29} \\ 0 & 0 & \beta_{33} & \beta_{34} & \beta_{35} & \beta_{36} & \beta_{37} & \beta_{38} & \beta_{39} \\ 0 & 0 & \beta_{43} & \beta_{44} & \beta_{45} & \beta_{46} & \beta_{47} & \beta_{48} & \beta_{49} \\ 0 & 0 & \beta_{53} & \beta_{54} & \beta_{55} & \beta_{56} & \beta_{57} & \beta_{58} & \beta_{59} \\ 0 & 0 & \beta_{63} & \beta_{64} & \beta_{65} & \beta_{66} & \beta_{67} & \beta_{68} & \beta_{69} \end{bmatrix}$.

iv. Note that $\mathbf{b}_3$ is not a linear combination of $\mathbf{b}_1$.
Then there is at least one non-zero entry in $\mathbf{b}_3$ from the second entry downwards.

Now, since B is a reduced row-echelon form, $\mathbf{b}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{c}_3$.

v. Note that $\mathbf{b}_4$ is not a linear combination of $\mathbf{b}_1, \mathbf{b}_3$.
Then there is at least one non-zero entry in $\mathbf{b}_3$ from the third entry downwards.

Now, since B is a reduced row-echelon form, $\mathbf{b}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{c}_4$.

vi. Recall that $\mathbf{c}_5 = \mathbf{c}_3 + \mathbf{c}_4$, $\mathbf{c}_6 = \mathbf{c}_1 - 2\mathbf{c}_3 - \mathbf{c}_4$, and $\mathbf{c}_7 = \mathbf{c}_1 + 3\mathbf{c}_3 + 2\mathbf{c}_4$.
Then, since B is row-equivalent to C, $\mathbf{b}_5 = \mathbf{b}_3 + \mathbf{b}_4$, $\mathbf{b}_6 = \mathbf{b}_1 - 2\mathbf{b}_3 - \mathbf{b}_4$, $\mathbf{b}_7 = \mathbf{b}_1 + 3\mathbf{b}_3 + 2\mathbf{b}_4$.
Since $\mathbf{b}_1 = \mathbf{c}_1$, $\mathbf{b}_3 = \mathbf{c}_3$ and $\mathbf{b}_4 = \mathbf{c}_4$, we have $\mathbf{b}_5 = \mathbf{c}_5$, $\mathbf{b}_6 = \mathbf{c}_6$ and $\mathbf{b}_7 = \mathbf{c}_7$ also.

So now $B = \begin{bmatrix} 1 & -2 & 0 & 0 & 0 & 1 & 1 & \beta_{18} & \beta_{19} \\ 0 & 0 & 1 & 0 & 1 & -2 & 3 & \beta_{28} & \beta_{29} \\ 0 & 0 & 0 & 1 & 1 & -1 & 2 & \beta_{38} & \beta_{39} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta_{48} & \beta_{49} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta_{58} & \beta_{59} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta_{68} & \beta_{69} \end{bmatrix}$.

vii. Note that $\mathbf{b}_8$ is not a linear combination of $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_4$.
Then there is at least one non-zero entry in $\mathbf{b}_8$ from the fourth entry downwards.

Now, since B is a reduced row-echelon form, $\mathbf{b}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{c}_8$.

viii. Recall that $\mathbf{c}_9 = -2\mathbf{c}_1 + 2\mathbf{c}_3 + 3\mathbf{c}_4 - 3\mathbf{c}_8$.
Then, since B is row-equivalent to C, $\mathbf{b}_9 = -2\mathbf{b}_1 + 2\mathbf{b}_3 + 3\mathbf{b}_4 - 3\mathbf{b}_8$.
Since $\mathbf{b}_1 = \mathbf{c}_1$, $\mathbf{b}_3 = \mathbf{c}_3$, $\mathbf{b}_4 = \mathbf{c}_4$ and $\mathbf{b}_8 = \mathbf{c}_8$, we have $\mathbf{b}_9 = \mathbf{c}_9$ also.

Hence $B = \begin{bmatrix} 1 & -2 & 0 & 0 & 0 & 1 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 & 1 & -2 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 & -1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = C$.