2.2 Row-echelon forms and reduced row-echelon forms.

0. *Assumed background.*

   - 2.1 *Systems of linear equations.*

   Preferred to have been prepared with.

   - 1.5 *Linear combinations.*
   - 1.6 *Linear dependence and linear independence.*

   Abstract. We introduce:—

   - the notion of row-echelon forms, and rank of row-echelon forms,
   - the notions of reduced row-echelon forms, leading ones, pivot columns and free columns.

1. **Definition. (Row-echelon form.)**

   Let C be a $(p \times q)$-matrix.

   We say that C is a **row-echelon form** *if and only if the statements below hold:*

   (1) *All rows consisting of only 0's are beneath the non-zero rows of C.*

   (2) *In each pair of adjacent non-zero rows, the first (or left-most) non-zero entry in the row above is always strictly to the left of that in the row below.*

   The number of non-zero rows in C is called the **rank of the row-echelon form** *C.*

   A column of C in which the first (or left-most) non-zero entry of some non-zero row is located is called a **pivot column of** *C.*

   A column of C which is not a pivot column is called a **free column of** *C.*

   Further convention and terminology.

   In such a row-echelon form C with, say, r non-zero rows (and hence r pivot columns):—

   - the pivot columns are labelled, from left to right, the d_1-th, d_2-th, d_3-th, ... d_r-th columns, and
   - the free columns are labelled, from left to right, the f_1-th, f_2-th, f_3-th, ..., f_{q-r}-th columns.

   Visualization.

   In such a row-echelon form C, the zeros to the left of the first non-zero entries of the various non-zero rows form something like a 'staircase' of zeros, and the first non-zero entries form something like step-edges of this 'staircase'.

   As there are r non-zero rows in total, C reads like:—

$$\left[\begin{array}{ccc|c|ccc|c|ccc|c|ccc|c|ccc}
0 & \cdots & 0 & \sharp_1 & \star & \cdots & \star & \star & \star & \cdots & \star & \star & \star & \cdots & \star & \star & \star & \cdots & \star \\ \hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \sharp_2 & \star & \cdots & \star & \star & \star & \cdots & \star & \star & \star & \cdots & \star \\ \hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \sharp_3 & \star & \cdots & \star & \star & \star & \cdots & \star \\ \hline
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & & \ddots & & \vdots & \vdots & & \vdots \\ \hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \sharp_r & \star & \cdots & \star \\ \hline
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0}
\end{array}\right]$$

   The symbol $\mathbf{0}$ stands for a column of zeros.

   The symbols $\sharp_1, \sharp_2, \sharp_3, \cdots, \sharp_r$ stand for the respective first non-zero entry in the non-zero rows. The columns in which these entries are located are the pivot columns of C: they are the d_1-th, d_2-th, d_3-th, ... d_r-th columns respectively.

   Remark. This version of definition is one of several (slightly) different (but equally reasonable) versions of definitions for the phrase *row-echelon form.*

2. **Example (1). (Matrices which are row-echelon forms, and those which are not.)**

   (a) These are row-echelon forms:

   i. $\begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

   ii. $\begin{bmatrix} 2 & 3 & 5 & 7 & 9 & 0 \\ 0 & 0 & 4 & 2 & 0 & 1 \\ 0 & 0 & 0 & 6 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix}$

   iii. $\left[\begin{array}{cccccc} 1 & 3 & 5 & 7 & 9 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

   iv. $\left[\begin{array}{cccccc} 1 & 3 & 5 & 7 & 9 & 0 \\ 0 & 0 & 4 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

   v. $\left[\begin{array}{c|ccccc} 0 & 1 & 3 & 5 & 7 & 9 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

   vi. $\left[\begin{array}{c|ccccc} 0 & 2 & 3 & 4 & 5 & 6 \\ 0 & 0 & 0 & 0 & 4 & 8 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(b) These are not row-echelon forms:

i. $\begin{bmatrix} 1 & 3 & 5 & 7 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ ii. $\begin{bmatrix} 0 & 1 & 3 & 5 & 7 & 9 \\ 1 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ iii. $\begin{bmatrix} 0 & 0 & 3 & 5 & 7 & 9 \\ 0 & 0 & 2 & 4 & 6 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3. **Definition. (Reduced row-echelon form.)**

Let C be a $(p \times q)$-matrix.

We say C is said to be a **reduced row-echelon form** *if and only if C is a row-echelon form and furthermore, the statements below hold:—*

(3) *In each non-zero row, the first (or left-most) non-zero entry is* 1. *Such an entry is called a* **leading one**.

(4) *Each leading one is the only non-zero entry in the column where it is located.*

Visualization.
Such a reduced row-echelon form C reads like:—

$$\left[\begin{array}{ccc|c|ccc|c|ccc|c|ccc|c|ccc}
0 & \cdots & 0 & 1 & \star & \cdots & \star & 0 & \star & \cdots & \star & 0 & \star & \cdots & \star & 0 & \star & \cdots & \star \\ \hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \star & \cdots & \star & 0 & \star & \cdots & \star & 0 & \star & \cdots & \star \\ \hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \star & \cdots & \star & 0 & \star & \cdots & \star \\ \hline
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & & \ddots & & \mathbf{0} & \vdots & & \vdots \\ \hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \star & \cdots & \star \\ \hline
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0}
\end{array}\right]$$

The zeros to the left of the leading ones of the various non-zero rows form something like a 'staircase' of zeros, and the leading ones form something like step-edges of this 'staircase'.

Remark. The rank of the reduced row-echelon form C is simultaneously

- the number of non-zero rows in C,
- the number of leading ones in C, and
- the number of pivot columns in C.

The pivot columns of C are those columns of C which contain leading ones of C.

The free columns of C are those columns of C which do not contain leading ones of C.

Further remark. Looking carefully at the 'visualization' of the reduced row-echelon form C (and also the concrete examples of reduced row-echelon forms), we may suspect immediately that in any reduced row-echelon form:—

(a) the pivot columns are linearly independent,

(b) each free column is a linear combination of the pivot columns strictly to its left, and

(c) the non-zero rows are linearly independent.

In fact, this is also the case in row-echelon forms. However, as we do not need this result at the moment, we will omit it for now.

4. **Example (2). (Row-echelon forms which are reduced row-echelon forms, and those which are not).**

(a) These are reduced row-echelon forms:

i. $\begin{bmatrix} 1 & 3 & 0 & 0 & 9 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ ii. $\left[\begin{array}{cccccc} 1 & 3 & 0 & 0 & 9 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$ iii. $\left[\begin{array}{c|ccccc} 0 & 1 & 3 & 0 & 0 & 9 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(b) These are not reduced row-echelon forms, despite being row-echelon forms:

i. $\begin{bmatrix} 1 & 3 & 0 & 0 & 9 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ ii. $\left[\begin{array}{cccccc} 1 & 3 & 0 & 0 & 9 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$ iii. $\left[\begin{array}{c|ccccc} 0 & 1 & 3 & 0 & 0 & 9 \\ 0 & 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 3 & 6 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$

Remark. Every reduced row-echelon form is a row-echelon form in the first place. If a matrix is not a row-echelon form, then it is definitely not a reduced row-echelon form.

5. **How to check whether a given matrix is a row-echelon form, or a reduced row-echelon form?**

   Suppose C is a matrix.

   Proceeding as described below, we check whether C is a row-echelon form, and whether C is a reduced row-echelon form.

   Step (1). Inspect C and ask:—

   Is it of the form $\left[\begin{array}{c|c} \mathcal{O} & \tilde{C} \\ \hline \mathcal{O} & \mathcal{O} \end{array}\right]$ *for some matrix* $\tilde{C}$ *whose* $(1,1)$*-th entry is non-zero and whose last row is non-zero?*

   (Here the $\mathcal{O}$'s stand for zero matrices of various sizes.)

   - If *no*, conclude that C is not a row-echelon form.
   - If *yes*, go to Step (2).

   Step (2). Inspect $\tilde{C}$ row by row, and ask, for each pair of consecutive rows:—

   Is the first non-zero entry in the row above strictly to the left of that in the row below it?

   - If *no* (for some pair of consecutive rows), conclude that C is not a row-echelon form.
   - If *yes* (for every pair of consecutive rows), conclude that C is a row echelon form (and $\tilde{C}$ is a row-echelon form), and go to Step (3).

   Step (3). Inspect the respective first non-zero entries in each row, and ask:—

   Are these entries all 1*'s?*

   - If no, conclude that C is not a reduced row-echelon form (and $\tilde{C}$ is a not a reduced row-echelon form).
   - If yes, go to Step (4).

   Step (4). Inspect the columns which contain these entries, and ask:—

   Is each of these 1*'s the only non-zero entry in the respective column in* $\tilde{C}$*?*

   - If *no* (for some such column), conclude that C is not a reduced row-echelon form (and $\tilde{C}$ is a not a reduced row-echelon form).
   - If *yes* (for every such column), conclude that C is a reduced row-echelon form (and $\tilde{C}$ is a reduced row-echelon form).

6. **Question.** Why are we interested in reduced row-echelon forms?

   Answer. We are interested in reduced row-echelon forms because it is easy to read off information about solutions of a system of linear equations when its augmented matrix representation is a reduced row-echelon form. To be more detailed:—

   - It is easy to read off from the matrix whether the system is consistent.
   - If it is consistent, it is easy to read off from the matrix all solutions of the system.
   - The matrix itself suggests how to present all solutions in a systematic and economic way, as linear combinations of a few linearly independent column vectors whose entries can be read off from the matrix.

7. **Example (3). (Motivation and illustration for the content of Theorem (1) and Theorem (2).)**

   (a) Let $C = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right]$.

   Note that C is a reduced row-echelon form, with:—

   - 1-st, 2-nd, 3-rd, 4-th columns as pivot columns, and
   - 5-th column as its (only) free column.

   C is the augmented matrix representation of the system $\mathcal{LS}(A, \mathbf{b})$, in which

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

When it is written out explicitly as simultaneous equations (for numbers) whose respective unknowns are x_1, x_2, x_3, x_4, it reads:—

$$\begin{cases} x_1 & & & & = & 1 \\ & x_2 & & & = & 2 \\ & & x_3 & & = & 3 \\ & & & x_4 & = & 4 \end{cases}$$

It follows that the system $\mathcal{LS}(A,\, \mathbf{b})$ has one and only solution, namely $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$.

Hence $\mathcal{LS}(A,\, \mathbf{b})$ is consistent.

(b) Let $C = \left[\begin{array}{ccccc|c} 1 & 3 & 0 & 0 & 9 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$.

Note that C is a reduced row-echelon form, with:—

- 1-st, 3-rd, 4-th, 6-th columns as pivot columns, and
- 2-nd, 5-th columns as free columns.

C is the augmented matrix representation of the system $\mathcal{LS}(A,\, \mathbf{b})$, in which

$$A = \begin{bmatrix} 1 & 3 & 0 & 0 & 9 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

When it is written out explicitly as simultaneous equations (for numbers) whose respective unknowns are x_1, x_2, x_3, x_4, x_5, the fourth equation reads:—

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 + 0 \cdot x_5 = 1,$$

from which no equality will be yielded no matter what numbers are substituted into the unknowns.
It follows that the system $\mathcal{LS}(A,\, \mathbf{b})$ is inconsistent.

(c) Let $C = \left[\begin{array}{ccccc|c} 1 & 3 & 0 & 0 & 9 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$.

Note that C is a reduced row-echelon form, with:—

- 1-st, 3-rd, 4-th columns as pivot columns, and
- 2-nd, 5-th, 6-th columns as free columns.

C is the augmented matrix representation of the system $\mathcal{LS}(A,\, \mathbf{b})$, in which

$$A = \begin{bmatrix} 1 & 3 & 0 & 0 & 9 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}.$$

i. We write out the system explicitly as simultaneous equations (for numbers) whose respective unknowns are x_1, x_2, x_3, x_4, x_5. It reads:—

$$\begin{cases} x_1 & + & 3x_2 & & & + & 9x_5 & = & 0 \\ & & & x_3 & & & & = & 1 \\ & & & & x_4 & + & 4x_5 & = & 2 \\ & & & & & & 0 & = & 0 \end{cases}$$

ii. From these equations, we see that a substitution of '$x_2 = 0,\ x_5 = 0$' and '$x_1 = 0,\ x_3 = 1,\ x_4 = 2$' yield equalities everywhere.

It follows that $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}$ is a solution of $\mathcal{LS}(A,\, \mathbf{b})$.

Hence $\mathcal{LS}(A,\, \mathbf{b})$ is consistent.

iii. We now prepare to read off all solutions of $\mathcal{LS}(A,\, \mathbf{b})$:—

- we first ignore the equation '$0 = 0$' (which provides no information), and

- we then rewrite the remaining equations as a collection of three simultaneous 'relations', which 'relate' each unknown with x_2, x_5 alone:—

$$(S) \quad \left\{ \begin{array}{lcccc} x_1 & = & & -3x_2 & -9x_5 \\ x_3 & = & 1 & & \\ x_4 & = & 2 & & -4x_5 \end{array} \right.$$

Now, recalling the definitions for vector equality, vector addition and scalar multiplication, we re-write (S) as:—

$$(S') \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -9 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

Remember (S') is $\mathcal{LS}(A, \mathbf{b})$ in disguise.

iv. We are now ready to read off all solutions of $\mathcal{LS}(A, \mathbf{b})$.
Note that (S') informs us of two things:—

A. If $\mathbf{t}$ is a solution of $\mathcal{LS}(A, \mathbf{b})$ then

$$\mathbf{t} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} + u \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v \begin{bmatrix} -9 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix} \quad \text{for some numbers } u, v.$$

B. If

$$\mathbf{t} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} + u \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v \begin{bmatrix} -9 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix} \quad \text{for some numbers } u, v,$$

then $\mathbf{t}$ is a solution of $\mathcal{LS}(A, \mathbf{b})$.

So it follows that a full description of all solutions of $\mathcal{LS}(A, \mathbf{b})$ is given by:—

- $\mathbf{t}$ *is a solution of* $\mathcal{LS}(A, \mathbf{b})$ *if and only if*

$$\textit{there are some numbers } u, v \textit{ such that } \mathbf{t} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} + u \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v \begin{bmatrix} -9 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix}.$$

v. We further observe that the column vectors constructed with the entries of the non-last free columns of C in this process, namely, $\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -9 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix}$, are linearly independent. Justification:—

- Let α_1, α_2 be real numbers. Suppose $\alpha_1 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} -9 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix} = \mathbf{0}_5$.

Then $\begin{bmatrix} -3\alpha_1 - 9\alpha_2 \\ \alpha_1 \\ 0 \\ -4\alpha_2 \\ \alpha_2 \end{bmatrix} = \alpha_1 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} -9 \\ 0 \\ 0 \\ -4 \\ 1 \end{bmatrix} = \mathbf{0}_5$.

By the definition of matrix equality, we have $\alpha_1 = 0$ and $\alpha_2 = 0$.

8. **Theorem (1).**

Let A be an $(m \times n)$-matrix, and $\mathbf{b}$ be a column vector with m entries.

Denote by C the augmented matrix representation of $\mathcal{LS}(A, \mathbf{b})$. (So $C = [\ A \mid \mathbf{b}\ ]$.)

Suppose C is a reduced row-echelon form with rank r.

Then the entries of A, $\mathbf{b}$ beneath the respective r-th rows are all 0.

Moreover, the statements below are logically equivalent:—

(a) *$\mathcal{LS}(A, \mathbf{b})$ is consistent.*

(b) *The last column of C is a free column.*

(c) *No row of C reads* $[\ 0 \quad 0 \quad \cdots \quad 0 \quad 0 \quad 1\ ]$.

9. **Theorem (2). (Full description of solutions for a consistent system with augmented matrix representation being a reduced row-echelon form.)**

Let A be an $(m \times n)$-matrix, and $\mathbf{b}$ be a column vector with m entries.

Denote by C the augmented matrix representation of $\mathcal{LS}(A, \mathbf{b})$. (So $C = [\ A \mid \mathbf{b}\ ]$.)

Suppose C is a reduced row-echelon form with r leading ones,

- *whose pivot columns, from left to right, are the d_1-th, d_2-th, ..., d_r-th columns, and*
- *whose free columns, from left to right, are the f_1-th, f_2-th, ..., f_{n-r}-th, f_{n+1-r}-th columns.*

Also suppose $d_1 = 1$.

Suppose $\mathcal{LS}(A, \mathbf{b})$ is consistent.

Then the statements below hold:—

(a) $r \leq n$.

(b) *The $f_{(n+1-r)}$-th column of C is the last column of C, namely, $\mathbf{b}$.*

(c) *Denote the top r entries of $\mathbf{b}$ by $b_1, b_2, \cdots, b_r$, from the top downwards.*
Denote by $\mathbf{p}$ the column vector with n entries in which:—
- *the d_1-th, d_2-th, ..., d_r-th entries are $b_1, b_2, \cdots, b_r$ respectively, and*
- *all other entries are 0.*

Then $\mathbf{p}$ is a (particular) solution of $\mathcal{LS}(A, \mathbf{b})$.

(d) *Suppose $r = n$.*
(So there is no free column in C other than the last column of C.)
Then $\mathbf{p}$ is the one and only one solution of $\mathcal{LS}(A, \mathbf{b})$.

(e) *Suppose $r < n$ (instead of supposing '$r = n$').*
(So some other column in C other than the last column of C is a free column.)
For each $\ell = 1, 2, \cdots, n-r$, denote the top r entries of the f_ℓ-th column by $\alpha_{1\ell}, \alpha_{2\ell}, \cdots, \alpha_{r\ell}$, from the top downwards.
Further denote by $\mathbf{q}_\ell$ the column vector with n entries, in which:—
- *the d_1-th, d_2-th, ..., d_r-th entries are $-\alpha_{1\ell}, -\alpha_{2\ell}, \cdots, -\alpha_{r\ell}$,*
- *the f_ℓ-th entry is 1, and*
- *all other entries are 0.*

Then the statements below hold:—

i. *The column vectors $\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_{n-r}$ are linearly independent.*

ii. *Suppose $\mathbf{t}$ is a column vector with n entries.*
Then $\mathbf{t}$ is a solution of $\mathcal{LS}(A, \mathbf{b})$ if and only if
there are some numbers $u_1, u_2, \cdots, u_{n-r}$ such that $\mathbf{t} = \mathbf{p} + u_1\mathbf{q}_1 + u_2\mathbf{q}_2 + \cdots + u_{n-r}\mathbf{q}_{n-r}$.

iii. *The system $\mathcal{LS}(A, \mathbf{b})$ has distinct solutions.*

Remark. We omit the argument for Theorem (1) and Theorem (2), as this is only a tedious exercise in book-keeping.

In terms of the symbols in the statement of Theorem (2), the conclusion of Theorem (2) informs us on how to read off all solutions of $\mathcal{LS}(A, \mathbf{b})$ in concrete situations:—

- With the respective unknowns $x_1, x_2, \cdots, x_n$ displayed explicitly, we may re-write the system as the collection of r simultaneous 'relations', relating each unknown with $x_{f_1}, x_{f_2}, \cdots, x_{n-r}$ alone:—

$$(S): \left\{ \begin{array}{lllllllllllll} x_{d_1} & = & b_1 & - & \alpha_{11}x_{f_1} & - & \alpha_{12}x_{f_2} & - & \alpha_{13}x_{f_3} & - & \cdots & - & \alpha_{1,n-r}x_{f_{n-r}} \\ x_{d_2} & = & b_2 & - & \alpha_{21}x_{f_1} & - & \alpha_{22}x_{f_2} & - & \alpha_{23}x_{f_3} & - & \cdots & - & \alpha_{2,n-r}x_{f_{n-r}} \\ x_{d_3} & = & b_3 & - & \alpha_{31}x_{f_1} & - & \alpha_{32}x_{f_2} & - & \alpha_{33}x_{f_3} & - & \cdots & - & \alpha_{3,n-r}x_{f_{n-r}} \\ & \vdots & & & & & & & & & & & \\ x_{d_r} & = & b_r & - & \alpha_{r1}x_{f_1} & - & \alpha_{r2}x_{f_2} & - & \alpha_{r3}x_{f_3} & - & \cdots & - & \alpha_{r,n-r}x_{f_{n-r}} \end{array} \right.$$

In fact $\alpha_{k\ell} = 0$ whenever $k > \ell$ (because C is a reduced row-echelon form). So even here there are a lot of 0's.

Now further using the definitions for vector equality, vector addition and scalar multiplication, we re-write (S) as:—

$$(S'):\quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{p} + x_{f_1}\mathbf{q}_1 + x_{f_2}\mathbf{q}_2 + x_{f_3}\mathbf{q}_3 + \cdots + x_{f_{n-r}}\mathbf{q}_{n-r}$$

It is the re-expression of $\mathcal{LS}(A,\, \mathbf{b})$ as (S') which tells us how to present all solutions of $\mathcal{LS}(A,\, \mathbf{b})$ in a systematic and economic way, as linear combinations of several linearly independent column vectors whose entries are read off from C.

10. Theorem (1) and Theorem (2) cover all systems of linear equations, whether homogeneous or non-homogeneous. However, for a homogeneous system, we only need to focus on its coefficient matrix.

Theorem (3).

Let A be an $(m \times n)$-matrix, and C be the augmented matrix representation of the homogeneous system $\mathcal{LS}(A,\, \mathbf{0}_m)$.

Suppose A is a reduced row-echelon form whose first column is a pivot column and which has r leading ones.

Then C is a reduced row-echelon form whose first column is a pivot column and which has r leading ones.

Moreover:—

(a) *$\mathcal{LS}(A,\, \mathbf{0}_m)$ is consistent, with $\mathbf{0}_n$ being a solution of the system.*

(b) *The inequality $r \leq n$ holds.*

(c) i. If $r = n$ then $\mathbf{0}_n$ is the one and only one solution of $\mathcal{LS}(A,\, \mathbf{0}_m)$.

ii. If $r < n$ then $\mathcal{LS}(A,\, \mathbf{b})$ has distinct solutions, and in particular, some non-trivial solution.

Proof of Theorem (3). Exercise. (How A and C relate with each other is a game of words on definitions. The rest of the result follows from Theorem (2).)

11. **Example (4). (Illustration on the content of Theorem (3).)**

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

Note that A is a reduced row-echelon form, with:—

- 1-st, 3-rd, 4-th, 5-th, 7-th columns as pivot columns, and
- 2-nd, 6-th, 8-th columns as free columns.

A is the coefficient matrix of the homogeneous system $\mathcal{LS}(A,\, \mathbf{0}_5)$.

We write out the system explicitly as simultaneous equations (for numbers) whose respective unknowns are $x_1, x_2, \cdots, x_8$. It reads:—

$$\left\{\begin{array}{lllllllllllll} x_1 & + & 2x_2 & & & & & + & 3x_6 & & - & 2x_8 & = 0 \\ & & & x_3 & & & & + & 5x_6 & & - & 3x_8 & = 0 \\ & & & & x_4 & & & + & 7x_6 & & + & 8x_8 & = 0 \\ & & & & & x_5 & & + & 9x_6 & & + & 6x_8 & = 0 \\ & & & & & & & & & x_7 & + & 4x_8 & = 0 \end{array}\right.$$

To see how to read off all solutions of $\mathcal{LS}(A,\, \mathbf{0}_5)$, we rewrite the equations as a collection of three simultaneous 'relations', which 'relate' each unknown with x_2, x_6, x_8 alone:—

$$\left\{\begin{array}{lllll} x_1 & = & -2x_2 & -3x_6 & +2x_8 \\ x_3 & = & & -5x_6 & +3x_8 \\ x_4 & = & & -7x_6 & -8x_8 \\ x_5 & = & & -9x_6 & -6x_8 \\ x_7 & = & & & -4x_8 \end{array}\right.$$

Now, using the definitions for vector equality, vector addition and scalar multiplication, we re-write these relations as:—

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} -3 \\ 0 \\ -5 \\ -7 \\ -9 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_8 \begin{bmatrix} 2 \\ 0 \\ 3 \\ -8 \\ -6 \\ 0 \\ 1 \\ -4 \end{bmatrix}$$

It follows that a full description of all solutions of $\mathcal{LS}(A,\, \mathbf{0}_5)$ is given by:—

- $\mathbf{t}$ *is a solution of* $\mathcal{LS}(A,\, \mathbf{0}_5)$ *if and only if*

$$\text{there are some numbers } u, v, w \text{ such that } \mathbf{t} = u\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v\begin{bmatrix} -3 \\ 0 \\ -5 \\ -7 \\ -9 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w\begin{bmatrix} 2 \\ 0 \\ 3 \\ -8 \\ -6 \\ 0 \\ 1 \\ -4 \end{bmatrix}.$$

We note that the column vectors constructed with the entries of the free columns of A in this process, namely, $\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ -5 \\ -7 \\ -9 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 3 \\ -8 \\ -6 \\ 0 \\ 1 \\ -4 \end{bmatrix}$ are linearly independent.

(b) Let $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Note that A is a reduced row-echelon form, with every column being a pivot column.

A is the coefficient matrix of the homogeneous system $\mathcal{LS}(A,\, \mathbf{0}_4)$.

We write out the system explicitly as simultaneous equations (for numbers) whose respective unknowns are x_1, x_2, x_3, x_4. It reads:—

$$\left\{\begin{array}{ccccccc} x_1 & & & & & = & 0 \\ & & x_2 & & & = & 0 \\ & & & x_3 & & = & 0 \\ & & & & x_4 & = & 0 \\ & & & & 0 & = & 0 \end{array}\right.$$

This tells us that the zero vector $\mathbf{0}_4$ is the one and only one solution of the system $\mathcal{LS}(A,\, \mathbf{0}_5)$.

The homogeneous system $\mathcal{LS}(A,\, \mathbf{0}_5)$ has no non-trivial solution.