4.8 Rank-Nullity Formulae.

0. *Assumed background.*
   - Whatever has been covered in Topics 1-3.
   - 4.5 *Dimension for subspaces of column vectors.*
   - 4.6 *Span, column space, row space, and minimal spanning set.*
   - 4.7 *Basis and dimension, null space, and the notion of subspace of a subspace.*

   Abstract. We introduce:—
   - the Rank-nullity Formulae,
   - various applications of the Rank-nullity Formulae, especially in systems of linear equations.

1. Recall the definitions for the notions of nullity, column rank, row rank, and some key results on them introduced earlier.

   Definition. (Nullity, column rank, row rank of an arbitrary matrix.)

   Let A be a $(p \times q)$-matrix with real entries.

   (a) *The* **null space of** A *is defined to be the subspace of* $\mathbb{R}^q$ *given by* $\{\ \mathbf{x} \in \mathbb{R}^q \mid A\mathbf{x} = \mathbf{0}_p\ \}$.
   It is denoted by $\mathcal{N}(A)$.
   The dimension of $\mathcal{N}(A)$ *is called the* **nullity of** A.

   (b) *Denote the columns of A, from left to right, by* $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_q$.
   The **column space** *of A is defined to be the subspace of* $\mathbb{R}^p$ *given by* $\mathsf{Span}\ (\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_q\})$.
   It is denoted by $\mathcal{C}(A)$.
   The dimension of $\mathcal{C}(A)$ *is called the* **column rank of** A.

   (c) *Write $B = A^t$, and denote the columns of B, from left to right, by* $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_p$.
   (So $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_p$ are the transposes of the respective rows of A from top to bottom.)
   The **row space** *of A is defined to be the subspace of* $\mathbb{R}^q$ *given by* $\mathsf{Span}\ (\{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_p\})$.
   It is denoted by $\mathcal{R}(A)$.
   The dimension of $\mathcal{R}(A)$ *is called the* **row rank of** A.

   Remark. By definition, $\mathcal{R}(A) = \mathcal{C}(A^t)$, and $\mathcal{R}(A^t) = \mathcal{C}(A)$.

   Theorem (♮).

   Let A be a $(p \times q)$-matrix with real entries.

   Suppose the rank of A is r. (So the rank of the reduced row-echelon form which is row-equivalent to A is r.)

   Then:—

   (1) $\dim(\mathcal{N}(A)) = q - r$.

   (2) $\dim(\mathcal{C}(A)) = \dim(\mathcal{R}(A)) = r$.

2. A seemingly trivial consequence of Theorem (♮) is the highly non-trivial result on matrices in general:—

   Theorem (1). (Rank-nullity Formulae.)

   Let A be a $(p \times q)$-matrix. (So the number of rows in A is p, and the number of columns in A is q.)

   Suppose that the rank of A is r.

   Then the equalities below hold:—

   (1) $\dim(\mathcal{N}(A)) + \dim(\mathcal{C}(A)) = q$.

   (2) $\dim(\mathcal{N}(A^t)) + \dim(\mathcal{R}(A)) = p$.

   Remark. The two equalities described in the conclusion of Theorem (1) are collectively known as the **Rank-nullity Formulae**.

   The second 'formulae' is an immediate consequence of the set equality $\mathcal{R}(A) = \mathcal{C}(A^t)$, and the application of Theorem (♮) onto the $(q \times p)$-matrix A^t.

3. **Example (1). (Illustrations of the Rank-nullity Formulae.)**

(a) Let $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & -1 & 0 \\ 3 & 4 & 4 & 3 \\ 2 & 2 & 1 & 1 \end{bmatrix}$, and write $B = A^t$.

Denote by A' the reduced row-echelon form which is row-equivalent to A.

Denote by B' the reduced row-echelon form which is row-equivalent to B.

i. Note that $A' = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

The pivot columns of A' are the 1-st, 2-nd, 3-rd columns. Hence the rank of A is 3.

A basis of $\mathcal{C}(A)$ is constituted by $\begin{bmatrix} 1 \\ 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 4 \\ 1 \end{bmatrix}$.

The dimension of $\mathcal{C}(A)$ is 3.

A basis of $\mathcal{N}(A)$ is constituted by $\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$.

The dimension of $\mathcal{N}(A)$ is 1.

As expected from theory, we have $\dim(\mathcal{N}(A)) + \dim(\mathcal{C}(A)) = 4$.

ii. Note that $B = \begin{bmatrix} 1 & 1 & 3 & 2 \\ 1 & 0 & 4 & 2 \\ 1 & -1 & 4 & 1 \\ 1 & 0 & 3 & 1 \end{bmatrix}$, and $B' = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

The pivot columns of B' are the 1-st, 2-nd, 3-rd columns. Hence the rank of B is 3.

A basis of $\mathcal{C}(B)$ is constituted by $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 4 \\ 3 \end{bmatrix}$.

The dimension of $\mathcal{C}(B)$ is 3.

A basis of $\mathcal{N}(B)$ is constituted by $\begin{bmatrix} 2 \\ -1 \\ -1 \\ 1 \end{bmatrix}$.

The dimension of $\mathcal{N}(B)$ is 1.

As expected from theory, we have $\dim(\mathcal{N}(A^t)) + \dim(\mathcal{R}(A)) = \dim(\mathcal{N}(A^t)) + \dim(\mathcal{C}(A^t)) = \dim(\mathcal{N}(B)) + \dim(\mathcal{C}(B)) = 4$.

(b) Let $A = \begin{bmatrix} 1 & 2 & 2 & 3 & 4 \\ 1 & 3 & 3 & 4 & 5 \\ 2 & 6 & 5 & 9 & 6 \end{bmatrix}$, and write $B = A^t$.

Denote by A' the reduced row-echelon form which is row-equivalent to A.

Denote by B' the reduced row-echelon form which is row-equivalent to B.

i. Note that $A' = \begin{bmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 & -3 \\ 0 & 0 & 1 & -1 & 4 \end{bmatrix}$.

The pivot columns of A' are the 1-st, 2-nd, 3-rd columns. Hence the rank of A is 3.

A basis of $\mathcal{C}(A)$ is constituted by $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$.

The dimension of $\mathcal{C}(A)$ is 3.

A basis of $\mathcal{N}(A)$ is constituted by $\begin{bmatrix} -1 \\ -2 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ -4 \\ 0 \\ 1 \end{bmatrix}$.

The dimension of $\mathcal{N}(A)$ is 2.

As expected from theory, we have $\dim(\mathcal{N}(A)) + \dim(\mathcal{C}(A)) = 5$.

ii. Note that $B = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 6 \\ 2 & 3 & 5 \\ 3 & 4 & 9 \\ 4 & 5 & 6 \end{bmatrix}$, and $B' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

The pivot columns of B' are the 1-st, 2-nd, 3-rd columns. Hence the rank of B is 3.

A basis of $\mathcal{C}(B)$ is constituted by $\begin{bmatrix}1\\2\\2\\3\\4\end{bmatrix}, \begin{bmatrix}1\\3\\3\\4\\5\end{bmatrix}, \begin{bmatrix}2\\6\\5\\9\\6\end{bmatrix}$.

The dimension of $\mathcal{C}(B)$ is 3.

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

The dimension of $\mathcal{N}(B)$ is 0.

As expected from theory, we have $\dim(\mathcal{N}(A^t)) + \dim(\mathcal{R}(A)) = \dim(\mathcal{N}(A^t)) + \dim(\mathcal{C}(A^t)) = \dim(\mathcal{N}(B)) + \dim(\mathcal{C}(B)) = 3$.

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

Denote by A' the reduced row-echelon form which is row-equivalent to A.

Denote by B' the reduced row-echelon form which is row-equivalent to B.

i. Note that $A' = \begin{bmatrix} 1 & -2 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

The pivot columns of A' are the 1-st, 3-rd, 4-th columns. Hence the rank of A is 3.

A basis of $\mathcal{C}(A)$ is constituted by $\begin{bmatrix}1\\0\\2\\3\end{bmatrix}, \begin{bmatrix}-1\\2\\-1\\-1\end{bmatrix}, \begin{bmatrix}1\\3\\3\\5\end{bmatrix}$.

The dimension of $\mathcal{C}(A)$ is 3.

A basis of $\mathcal{N}(A)$ is constituted by $\begin{bmatrix}2\\1\\0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\-1\\-1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\2\\1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\-3\\-2\\0\\0\\1\end{bmatrix}$.

The dimension of $\mathcal{N}(A)$ is 4.

As expected from theory, we have $\dim(\mathcal{N}(A)) + \dim(\mathcal{C}(A)) = 7$.

ii. Note that $B = \begin{bmatrix} 1 & 0 & 2 & 3 \\ -2 & 0 & -4 & -6 \\ -1 & 2 & -1 & -1 \\ 1 & 3 & 3 & 5 \\ 0 & 5 & 2 & 4 \\ 2 & -7 & 1 & 0 \\ 0 & 12 & 5 & 10 \end{bmatrix}$, $B' = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

The pivot columns of B' are the 1-st, 2-nd, 3-rd columns. Hence the rank of B is 3.

A basis of $\mathcal{C}(B)$ is constituted by $\begin{bmatrix}1\\-2\\-1\\1\\0\\2\\0\end{bmatrix}, \begin{bmatrix}0\\0\\2\\3\\5\\-7\\12\end{bmatrix}, \begin{bmatrix}2\\-4\\-1\\3\\2\\1\\5\end{bmatrix}$.

The dimension of $\mathcal{C}(B)$ is 3.

A basis of $\mathcal{N}(B)$ is constituted by $\begin{bmatrix}1\\0\\-2\\1\end{bmatrix}$.

The dimension of $\mathcal{N}(B)$ is 1.

As expected from theory, we have $\dim(\mathcal{N}(A^t)) + \dim(\mathcal{R}(A)) = \dim(\mathcal{N}(A^t)) + \dim(\mathcal{C}(A^t)) = \dim(\mathcal{N}(B)) + \dim(\mathcal{C}(B)) = 4$.

4. Recall the 'uniqueness' of $\{\mathbf{0}_n\}$ and $\mathbb{R}^n$ as 'extreme' subspaces of $\mathbb{R}^n$:—

Theorem ($\star$).

$\{\mathbf{0}_n\}$ *is the one and only one* 0*-dimensional subspace of* $\mathbb{R}^n$*, and* $\mathbb{R}^n$ *is the one and only one* n*-dimensional subspace of* $\mathbb{R}^n$*.*

We are now going to apply Theorem ($\natural$), Theorem ($\star$) and the Rank-nullity Formulae to deduce a highly non-trivial result about systems of linear equations (which has wide applications in other disciplines).

Theorem (2).

Suppose A is a $(p \times q)$-matrix with real entries. Then the statements below hold:—

($\sharp$) $\mathcal{C}(A) = \mathbb{R}^p$ *if and only if* $\mathcal{N}(A^t) = \{\mathbf{0}_p\}$.

($\sharp^*$) $\mathcal{R}(A) = \mathbb{R}^q$ *if and only if* $\mathcal{N}(A) = \{\mathbf{0}_q\}$.

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

   Suppose A is a $(p \times q)$-matrix with real entries.

   (a) [We want to verify ($\sharp$): $\mathcal{C}(A) = \mathbb{R}^p$ *if and only if* $\mathcal{N}(A^t) = \{\mathbf{0}_p\}$.]

   Preparation.

   By Theorem ($\natural$), $\dim(\mathcal{R}(A)) = \dim(\mathcal{C}(A))$.

   By the Rank-nullity Formula, $\dim(\mathcal{N}(A^t)) + \dim(\mathcal{R}(A)) = p$.

   Then $\dim(\mathcal{N}(A^t)) + \dim(\mathcal{C}(A)) = \dim(\mathcal{N}(A^t)) + \dim(\mathcal{R}(A)) = p$.

   Therefore $\dim(\mathcal{C}(A)) = p - \dim(\mathcal{N}(A^t))$. —— ($\triangle$)

   i. Suppose $\mathcal{C}(A) = \mathbb{R}^p$. Then by Theorem ($\star$), we have $\dim(\mathcal{C}(A)) = p$.
   Therefore, by ($\triangle$), we have $\dim(\mathcal{N}(A^t)) = 0$. Hence, by Theorem ($\star$), we have $\mathcal{N}(A^t) = \{\mathbf{0}_p\}$.

   ii. Suppose $\mathcal{N}(A^t) = \{\mathbf{0}_p\}$. Then by Theorem ($\star$), we have $\dim(\mathcal{N}(A^t)) = 0$.
   Therefore, by ($\triangle$), we have $\dim(\mathcal{C}(A)) = p$. Hence, by Theorem ($\star$), we have $\mathcal{C}(A) = \mathbb{R}^p$.

   (b) Define $B = A^t$. Note that B is a $(q \times p)$-matrix with real entries. Repeating the argument above, we deduce that $\mathcal{C}(B) = \mathbb{R}^q$ if and only if $\mathcal{N}(B^t) = \{\mathbf{0}_q\}$.

   Note that $\mathcal{C}(B) = \mathcal{C}(A^t) = \mathcal{R}(A)$ and $B^t = (A^t)^t = A$.

   Hence $\mathcal{R}(A) = \mathbb{R}^q$ if and only if $\mathcal{N}(A) = \{\mathbf{0}_q\}$.

6. We may explicitly re-formulate the content of Theorem (2) in the language of systems of linear equations:—

   Theorem (3). (Corollary to Theorem (2).)

   Suppose A is a $(p \times q)$-matrix with real entries. Then the results ($\sharp$), ($\sharp^*$) hold:—

   ($\sharp$) *The statements ($\sharp_1$), ($\sharp_2$) are logically equivalent:—*

   ($\sharp_1$) *For any* $\mathbf{b} \in \mathbb{R}^p$, *the system* $\mathcal{LS}(A, \mathbf{b})$ *is consistent.*

   ($\sharp_2$) *The homogeneous system* $\mathcal{LS}(A^t, \mathbf{0}_q)$ *has no non-trivial solution in* $\mathbb{R}^p$.

   ($\sharp^*$) *The statements ($\sharp_1^*$), ($\sharp_2^*$) are logically equivalent:—*

   ($\sharp_1^*$) *For any* $\mathbf{c} \in \mathbb{R}^q$, *the system* $\mathcal{LS}(A^t, \mathbf{c})$ *is consistent.*

   ($\sharp_2^*$) *The homogeneous system* $\mathcal{LS}(A, \mathbf{0}_p)$ *has no non-trivial solution in* $\mathbb{R}^q$.

   Remark. When A is a square matrix (and hence $p = q$), we know from the theory of invertibility that all four statements ($\sharp_1$), ($\sharp_2$), ($\sharp_1^*$), ($\sharp_2^*$) will be logical equivalent to each other. So what we have discovered here is a refinement of earlier results, through the use of more advanced theoretical machinery.

7. With a purely logical consideration, we also have the result below:—

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

   Suppose A is a $(p \times q)$-matrix with real entries. Then the results ($\widetilde{\sharp}$), ($\widetilde{\sharp^*}$) hold:—

   ($\widetilde{\sharp}$) *The statements ($\sim\sharp_1$), ($\sim\sharp_2$) are logically equivalent:—*

   ($\sim\sharp_1$) *There is some* $\mathbf{b} \in \mathbb{R}^p$ *such that the system* $\mathcal{LS}(A, \mathbf{b})$ *is inconsistent.*

   ($\sim\sharp_2$) *The homogeneous system* $\mathcal{LS}(A^t, \mathbf{0}_q)$ *has some non-trivial solution in* $\mathbb{R}^p$.

   ($\widetilde{\sharp^*}$) *The statements ($\sim\sharp_1^*$), ($\sim\sharp_2^*$) are logically equivalent:—*

   ($\sim\sharp_1^*$) *There is some* $\mathbf{c} \in \mathbb{R}^q$ *such that the system* $\mathcal{LS}(A^t, \mathbf{c})$ *is inconsistent.*

   ($\sim\sharp_2^*$) *The homogeneous system* $\mathcal{LS}(A, \mathbf{0}_p)$ *has some non-trivial solution in* $\mathbb{R}^q$.

8. **Example (2). (Illustration of the content of Theorem (3) and Theorem (4).)**

(a) Let

$$A = \begin{bmatrix} 1 & 2 & -5 & 15 \\ -1 & -1 & 3 & -9 \\ 3 & 4 & -10 & 31 \\ 2 & 3 & -8 & 25 \\ 1 & 3 & -4 & 13 \end{bmatrix}.$$

The reduced row-echelon form A' which is row-equivalent to A is given by

$$A' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Therefore $\mathcal{LS}(A,\, \mathbf{0}_5)$ has no non-trivial solution in $\mathbb{R}^4$.

It follows that for any $\mathbf{c} \in \mathbb{R}^4$, the system $\mathcal{LS}(A^t,\, \mathbf{c})$ is consistent.

So no matter which (real) values c_1, c_2, c_3, c_4 take, the system

$$\left\{ \begin{array}{rcrcrcrcrcl} x_1 & - & x_2 & + & 3x_3 & + & 2x_4 & + & x_5 & = & c_1 \\ 2x_1 & - & x_2 & + & 4x_3 & + & 3x_4 & + & 3x_5 & = & c_2 \\ -5x_1 & + & 3x_2 & - & 10x_3 & - & 8x_4 & - & 4x_5 & = & c_3 \\ 15x_1 & - & 9x_2 & + & 31x_3 & + & 25x_4 & + & 13x_5 & = & c_4 \end{array} \right.$$

is consistent.

(b) Let

$$A = \begin{bmatrix} 1 & 3 & 1 & -2 & 1 \\ 1 & 3 & 2 & -3 & -3 \\ 2 & 6 & 1 & -2 & 10 \\ -1 & -3 & -3 & 1 & -5 \\ 0 & 0 & 1 & -1 & -4 \\ 1 & 3 & -2 & -1 & 5 \end{bmatrix}.$$

The reduced row-echelon form A' which is row-equivalent to A is given by

$$A' = \begin{bmatrix} 1 & 3 & 0 & 0 & 9 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Therefore $\mathcal{LS}(A,\, \mathbf{0}_6)$ has some non-trivial solution in $\mathbb{R}^5$.

It follows that there is some $\mathbf{c} \in \mathbb{R}^6$ for which the system $\mathcal{LS}(A^t,\, \mathbf{c})$ is inconsistent.

(It will take further work to indeed pin down some such concrete $\mathbf{c} \in \mathbb{R}^6$.)

9. We can also give a re-interpretation of the content of Theorem (2) in terms of the notion of linear combination and linear independence.

Theorem (5). (Corollary to Theorem (2).)

Suppose A is a $(p \times q)$-matrix with real entries. Then the results $(\widehat{\sharp})$, $(\widehat{\sharp^*})$ hold:—

$(\widehat{\sharp})$ *The statements $(\widehat{\sharp_1})$, $(\widehat{\sharp_2})$ are logically equivalent:—*

- $(\widehat{\sharp_1})$ *Every column vector with p real entries is a linear combination of the columns of A over the reals.*
- $(\widehat{\sharp_2})$ *The rows of A are linearly independent over the reals.*

$(\widehat{\sharp^*})$ *The statements $(\widehat{\sharp_1^*})$, $(\widehat{\sharp_2^*})$ are logically equivalent:—*

- $(\widehat{\sharp_1^*})$ *Every row vector with q real entries is a linear combination of the rows of A over the reals.*
- $(\widehat{\sharp_2^*})$ *The columns of A are linearly independent over the reals.*

Remark. We also have a counterpart to Theorem (5), obtained with a purely logical consideration, that tells us what exactly will happen when the rows/columns of A are linearly dependent over the reals. Formulate it as an exercise.

10. We may wonder what can be said of a matrix whose column space and null space are not 'extreme' in the sense of Theorem ($\star$).

Theorem (6).

Suppose A is a $(p \times q)$-matrix with real entries.

Then the inequalities below hold:—

(1) $\dim(\mathcal{C}(A)) \leq \min(p, q)$, *(and* $\dim(\mathcal{R}(A)) \leq \min(p, q)$*).*

(2) $\dim(\mathcal{N}(A)) \geq \max(0, q - p)$.

(3) $\dim(\mathcal{N}(A^t)) \geq \max(0, p - q)$.

Remark. In the inequality (2), what is interesting is the scenario in which the strict inequality $p < q$ holds.

In this situation, the inequality '$\dim(\mathcal{N}(A)) \geq \max(0, q-p)$' is a very 'compact' way of presenting the following phenomenon about the homogeneous system $\mathcal{LS}(A, \mathbf{0}_p)$:—

- *When the number of unknowns in $\mathcal{LS}(A, \mathbf{0}_p)$ exceeds the number of equations in $\mathcal{LS}(A, \mathbf{0}_p)$ by, say, k, then the null space of A will have a basis which has at least k column vectors with q entries.*

11. **Example (3). (Illustration of the content of Theorem (6).)**

Let

$$A = \begin{bmatrix} 1 & 3 & 1 & -2 & 1 \\ 1 & 3 & 2 & -3 & -3 \\ 2 & 6 & 1 & -2 & 10 \\ -1 & -3 & -3 & 1 & -5 \end{bmatrix}.$$

Theorem (6) informs us that $\dim(\mathcal{N}(A)) \geq \max(0, 5-4) = \max(0, 1) = 1$.

Now we indeed compute $\dim(\mathcal{N}(A))$.

Note that the reduced row-echelon form A' which is row-equivalent to A is given by

$$A' = \begin{bmatrix} 1 & 3 & 0 & 0 & 9 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We have $\mathcal{N}(A) = \mathcal{N}(A')$, and hence $\dim(\mathcal{N}(A)) = \dim(\mathcal{N}(A')) = 2$ in fact.

12. **Proof of Theorem (6).**

Suppose A is a $(p \times q)$-matrix with real entries.

(1) Denote by A' the reduced row-echelon form row-equivalent to A.
Again recall that the dimension of $\mathcal{C}(A)$ is the rank r of A'.
As r is the number of non-zero rows in A', it is of value at most p.
Also, as r is the number simultaneously the number of pivot columns in A', it is of value at most q.
Hence r is of value at most the minimum between p, q.

(2) By the Rank-nullity formulae, we have $\dim(\mathcal{C}(A)) = q - \dim(\mathcal{N}(A))$.
Since $\dim(\mathcal{C}(A)) \leq p$, we have $p \geq q - \dim(\mathcal{N}(A))$.
Then $\dim(\mathcal{N}(A)) \geq q - p$.
Also note that $\dim(\mathcal{N}(A)) \geq 0$.
Therefore $\dim(\mathcal{N}(A)) \geq \max(0, q-p)$.

(3) Interchanging the respective roles of A, A^t, and of p, q in the argument above, we deduce that $\dim(\mathcal{C}(A^t)) \leq \min(q, p)$, and $\dim(\mathcal{N}(A^t)) \geq \max(0, p - q)$.

13. We now recall the inequality result for dimensions of 'comparable' subspaces, which is labeled Theorem ($\star\star$) here.

Theorem ($\star\star$).

Let $\mathcal{V}, \mathcal{W}$ be subspaces of $\mathbb{R}^q$ over the reals.

Suppose $\mathcal{V}$ is a subspace of $\mathcal{W}$ over the reals.

Then $\dim(\mathcal{V}) \leq \dim(\mathcal{W})$.

Moreover, equality holds if and only if $\mathcal{V} = \mathcal{W}$.

14. We are going to apply Theorem ($\natural$), Theorem ($\star\star$), the Rank-nullity Formulae, and Lemma (7), to deduce some formulae that relate the dimensions of null space, column space and row space for products of matrices with that for the individual matrices that make the products.

Lemma (7).

Suppose A is a $(p \times q)$-matrix with real entries, and B is a $(q \times s)$-matrix with real entries.

Then $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(AB)$ over the reals. Moreover, $\dim(\mathcal{N}(B)) \leq \dim(\mathcal{N}(AB))$.

Proof of Lemma (7).

Suppose A is a $(p \times q)$-matrix with real entries, and B is a $(q \times s)$-matrix with real entries.

By definition, AB is an $(p \times s)$-matrix. Note that $\mathcal{N}(B)$, $\mathcal{N}(AB)$ are both subspaces of $\mathbb{R}^s$ over the reals.

[We want to verify that $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(AB)$ over the reals.
This amounts to verifying the statement '*for any* $\mathbf{v} \in \mathbb{R}^s$, *if* $\mathbf{v} \in \mathcal{N}(B)$ *then* $\mathbf{v} \in \mathcal{N}(AB)$'.]

Pick any vector $\mathbf{v} \in \mathbb{R}^s$. Suppose $\mathbf{v} \in \mathcal{N}(B)$. Then by definition, $B\mathbf{v} = \mathbf{0}_q$.

We have $(AB)\mathbf{v} = A(B\mathbf{v}) = A\mathbf{0}_q = \mathbf{0}_p$. Then by definition $\mathbf{v} \in \mathcal{N}(AB)$.

It follows that $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(AB)$. By Theorem ($\star\star$), $\dim(\mathcal{N}(B)) \leq \dim(\mathcal{N}(AB))$ also.

15. **Theorem (8). (Upper bound of column rank of product of matrices.)**

Suppose A is a $(p \times q)$-matrix with real entries, and B is a $(q \times s)$-matrix with real entries.

Then the inequalities below hold:

(a) $\dim(\mathcal{C}(AB)) \leq \dim(\mathcal{C}(B))$.

(b) $\dim(\mathcal{C}(AB)) \leq \dim(\mathcal{C}(A))$.

(c) $\dim(\mathcal{C}(AB)) \leq \min(\dim(\mathcal{C}(A)), \dim(\mathcal{C}(B)))$.

16. An immediate consequence of the last part in Theorem (8) is the non-trivial result that relates the rank of the product of two matrices with the respective ranks of the individual matrices that give the product. This is not obvious if we are only looking at the reduced row-echelon forms row-equivalent to the respective matrices.

Theorem (9). (Corollary to Theorem (8).)

Suppose A is a $(p \times q)$-matrix with real entries, of rank r_A, and B is a $(q \times s)$-matrix with real entries, of rank r_B.

Then the rank of AB is of value at most $\min(r_A, r_B)$.

17. **Example (4). (Illustration on the content of Theorem (9).)**

Let $A = \begin{bmatrix} 1 & 3 & 1 & -2 & 1 \\ 1 & 3 & 2 & -3 & -3 \\ 2 & 6 & 1 & -2 & 10 \\ -1 & -3 & -3 & 1 & -5 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & 2 & 2 & 0 & 1 \\ 2 & 3 & 6 & 5 & 1 & 3 \\ 3 & 4 & 8 & 7 & 1 & 4 \\ 1 & 2 & 4 & 3 & 1 & 2 \\ 0 & 1 & 2 & 1 & 1 & 1 \end{bmatrix}$.

The reduced row-echelon form A' which is row-equivalent to A is given by

$$A' = \begin{bmatrix} 1 & 3 & 0 & 0 & 9 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence the rank of A is 3.

The reduced row-echelon form B' which is row-equivalent to B is given by

$$B' = \begin{bmatrix} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence the rank of B is 2.

Note that

$$AB = \begin{bmatrix} 8 & 11 & 22 & 19 & 3 & 11 \\ 10 & 9 & 18 & 19 & -1 & 9 \\ 15 & 30 & 60 & 45 & 15 & 30 \\ -15 & -25 & -50 & -40 & -10 & -25 \end{bmatrix}$$

and the reduced row-echelon form C' which is row-equivalent to AB is given by

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

Hence the rank of AB is 2, which is no greater than the respective ranks of A, B.

18. **Proof of Theorem (8).**

Suppose A is a $(p \times q)$-matrix with real entries, and B is a $(q \times s)$-matrix with real entries.

(a) By the Rank-nullity Formula, we have $\dim(\mathcal{N}(B)) + \dim(\mathcal{C}(B)) = s$, and $\dim(\mathcal{N}(AB)) + \dim(\mathcal{C}(AB)) = s$.
By Lemma (7), $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(AB)$, and the inequality $\dim(\mathcal{N}(B)) \leq \dim(\mathcal{N}(AB))$ holds.
Then $\dim(\mathcal{C}(AB)) = s - \dim(\mathcal{N}(AB)) \leq s - \dim(\mathcal{N}(B)) = \dim(\mathcal{C}(B))$.

(b) Note that $B^t A^t = (AB)^t$.
By Lemma (7), $\mathcal{N}(A^t)$ is a subspace of $\mathcal{N}((AB)^t)$, and the inequality $\dim(\mathcal{N}(A^t)) \leq \dim(\mathcal{N}((AB)^t))$ holds.
By the Rank-nullity Formula, we have $\dim(\mathcal{N}(A^t)) + \dim(\mathcal{R}(A)) = p$ and $\dim(\mathcal{N}((AB)^t)) + \dim(\mathcal{R}(AB)) = p$.
Then $\dim(\mathcal{R}(AB)) = p - \dim(\mathcal{N}((AB)^t)) \leq p - \dim(\mathcal{N}(A^t)) = \dim(\mathcal{R}(A))$.
Recall from Theorem (♮) that the equalities '$\dim(\mathcal{C}(AB)) = \dim(\mathcal{R}(AB))$', '$\dim(\mathcal{R}(A)) = \dim(\mathcal{C}(A))$' hold.
Then $\dim(\mathcal{C}(AB)) \leq \dim(\mathcal{C}(A))$.

(c) We have already verified the inequalities '$\dim(\mathcal{C}(AB)) \leq \dim(\mathcal{C}(B))$' and '$\dim(\mathcal{C}(AB)) \leq \dim(\mathcal{C}(A))$'.
It follows that $\dim(\mathcal{C}(AB)) \leq \min(\dim(\mathcal{C}(A)), \dim(\mathcal{C}(B)))$.

19. Earlier we have introduced the result below, which is labeled Theorem ($\star\star\star$).

Theorem ($\star\star\star$)**.**

Let A be a $(p \times q)$-matrix with real entries.

Suppose $\widehat{A}$ is a matrix with q columns whose rows are amongst the rows of A.
Then:—

(a) *$\mathcal{N}(A)$ is a subspace of $\mathcal{N}\left(\widehat{A}\right)$ over the reals.*

Moreover, $\dim(\mathcal{N}(A)) \leq \dim\left(\mathcal{N}\left(\widehat{A}\right)\right)$.

(b) *$\mathcal{R}\left(\widehat{A}\right)$ is a subspace of $\mathcal{R}(A)$ over the reals.*

Moreover, $\dim\left(\mathcal{R}\left(\widehat{A}\right)\right) \leq \dim(\mathcal{R}(A))$.

Equality holds if and only if each row of A is a linear combination of the rows of $\widehat{A}$.

20. It is tempting to insert at the end of part (a) of the conclusion of Theorem ($\star\star\star$) the words below:—

'*Equality holds if and only if each row of A is a linear combination of the rows of $\widehat{A}$.*'

It turns out that this 'upgrading' of Theorem ($\star\star\star$) is indeed correct. But this is a highly non-trivial matter, as it involves an application of the Rank-nullity Formulae.

Theorem (10)**.**

Let A be a $(p \times q)$-matrix with real entries.

Suppose $\widehat{A}$ is a matrix with q columns whose rows are amongst the rows of A.
Then:—

(a) i. *$\mathcal{N}(A)$ is a subspace of $\mathcal{N}\left(\widehat{A}\right)$ over the reals, and the inequality* $\dim(\mathcal{N}(A)) \leq \dim\left(\mathcal{N}\left(\widehat{A}\right)\right)$ *holds.*

ii. *$\mathcal{R}\left(\widehat{A}\right)$ is a subspace of $\mathcal{R}(A)$ over the reals, and the inequality* $\dim\left(\mathcal{R}\left(\widehat{A}\right)\right) \leq \dim(\mathcal{R}(A))$ *holds.*

(b) *The statements below are logically equivalent:—*

(b1) *Each row of A is a linear combination of the rows of $\widehat{A}$.*

(b2) $\dim(\mathcal{N}(A)) = \dim\left(\mathcal{N}\left(\widehat{A}\right)\right)$.

(b3) $\dim(\mathcal{R}(A)) = \dim\left(\mathcal{R}\left(\widehat{A}\right)\right)$.

(b4) $\mathcal{N}(A) = \mathcal{N}\left(\widehat{A}\right)$.

(b5) $\mathcal{R}(A) = \mathcal{R}\left(\widehat{A}\right)$.

21. **Comment on the argument for Theorem (10).**

Almost everything in Theorem (10) is a re-packaging of Theorem ($\star\star\star$). The only question is whether the five statements (b1)-(b5) stated in part (b) of its conclusion are indeed logically equivalent.

Theorem ($\star\star$) and Theorem ($\star\star\star$) in fact inform us that:—

- the statements (b2), (b4) are logically equivalent,
- the statements (b3), (b5) are logically equivalent, and
- the statements (b1), (b3) are logically equivalent.

So it suffices for us to simply deduce the logical equivalence of, say, the statements (b2), (b3).

22. **Proof of (the logical equivalence of the statements (b2), (b3) in the conclusion in) Theorem (10).**

Let A be a $(p \times q)$-matrix with real entries.

Suppose $\widehat{A}$ is a matrix with q columns whose rows are amongst the rows of A.

We verify that the statements (b2), (b3) are logically equivalent:—

(b2) $\dim(\mathcal{N}(A)) = \dim\left(\mathcal{N}\left(\widehat{A}\right)\right)$.

(b3) $\dim(\mathcal{R}(A)) = \dim\left(\mathcal{R}\left(\widehat{A}\right)\right)$.

By Theorem ($\natural$) and the Rank-nullity Formulae, we have

$$\begin{aligned}
\dim\left(\mathcal{N}\left(\widehat{A}\right)\right) + \dim\left(\mathcal{R}\left(\widehat{A}\right)\right) &= \dim\left(\mathcal{N}\left(\widehat{A}\right)\right) + \dim\left(\mathcal{C}\left(\widehat{A}\right)\right) \\
&= q \\
&= \dim(\mathcal{N}(A)) + \dim(\mathcal{C}(A)) \\
&= \dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A)).
\end{aligned}$$

It follows from this chain of equalities that $\dim(\mathcal{N}(A)) = \dim\left(\mathcal{N}\left(\widehat{A}\right)\right)$ if and only if $\dim(\mathcal{R}(A)) = \dim\left(\mathcal{R}\left(\widehat{A}\right)\right)$.

23. **Interpretation of part (b) in the conclusion of Theorem (10).**

The logical equivalence between (b1) and (b2) is a mathematically precise formulation of the 'fact' below about the homogeneous system $\mathcal{LS}(A,\, \mathbf{0}_p)$:—

Denote the rows of A, from top to bottom, by $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p$.

Suppose we can identify r linearly independent row vectors amognst $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p$, say, $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \cdots, \mathbf{a}_{k_r}$, of which every other of the $\mathbf{a}_j$'s is a linear combination.

In other words, we have obtained a basis for $\mathcal{R}(A)$ over the reals constituted by $\mathbf{a}_{k_1}{}^t, \mathbf{a}_{k_2}{}^t, \cdots, \mathbf{a}_{k_r}{}^t$.

Then, when we solve the system $\mathcal{LS}(A,\, \mathbf{0}_p)$, we may

- 'safely' ignore those equations in $\mathcal{LS}(A,\, \mathbf{0}_p)$ not amongst $\mathcal{LS}(\mathbf{a}_{k_1},\, \mathbf{0}_1)$, $\mathcal{LS}(\mathbf{a}_{k_2},\, \mathbf{0}_1)$, ..., $\mathcal{LS}(\mathbf{a}_{k_r},\, \mathbf{0}_1)$, and
- simply focus on solving the homogeneous system $\mathcal{LS}\left(\widehat{A},\, \mathbf{0}_r\right)$, in which $\widehat{A}$ is the $(r \times q)$-matrix whose rows, from top to bottom, are given by $\mathbf{a}_{k_1}, \mathbf{a}_{k_2}, \cdots, \mathbf{a}_{k_r}$.

Since $\mathcal{N}(A) = \mathcal{N}\left(\widehat{A}\right)$, a full description of all solutions of $\mathcal{LS}\left(\widehat{A},\, \mathbf{0}_r\right)$ is the same as a full description of all solutions of $\mathcal{LS}(A,\, \mathbf{0}_p)$.

(This is in fact the same essence of the application of Gaussian elimination in solving $\mathcal{LS}(A,\, \mathbf{0}_p)$. Gaussian elimination in fact takes the above approach 'one step further'. Instead of identifying a basis for $\mathcal{R}(A)$ over the reals from amongst $\mathbf{a}_1{}^t, \mathbf{a}_2{}^t, \cdots, \mathbf{a}_p{}^t$ only, we just work out a basis for $\mathcal{R}(A)$ over the reals which are given by the respective transposes of the non-zero rows of the reduced row-echelon form A' which is row-equivalent to A.)