3.3 Various necessary and sufficient conditions for invertibility.

0. *Assumed background.*

   - What has been covered in Topics 1-2.
   - 3.1 *Invertible matrices.*
   - 3.2 *Invertibility and row operations.*

   Abstract. We introduce:—

   - various re-formulations of invertibility, and the reasons behind.

1. Earlier we have introduced a result about necessary and sufficient conditions for invertibility in terms of row operations. It is in fact a 'baby version' of a more extensive result, which is Theorem (1) below.

   Theorem (1).

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

   (a) *A is invertible.*

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

   (c) *A is a product of $(p \times p)$-row-operation matrices.*

   (d) *A has a left inverse.*

   (e) *A has a right inverse.*

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

   (g) *For any column vector $\mathbf{b}$ with p entries, the system $\mathcal{LS}(A, \mathbf{b})$ is consistent.*

   Now suppose any one of the above holds (and hence all hold). Then the statements below hold:—

   (α) *$[\ A \mid I_p\ ]$ is row-equivalent to $[\ I_p \mid A^{-1}\ ]$.*

   (β) *I_p is the only reduced row-echelon form which is row-equivalent to A.*

   (γ) *For any column vector $\mathbf{b}$ with p entries, the system $\mathcal{LS}(A, \mathbf{b})$ has a unique solution, namely $A^{-1}\mathbf{b}$.*

2. **Comments on Theorem (1).**

   (a) Theorem (1) incorporates the claim ($\star$) we have made earlier but are yet to justify:—

   ($\star$) *Any given square matrix has both left and right inverses, or neither.*

   By proving Theorem (1), we can justify the claim ($\star$)

   (b) The full argument for Theorem (1) is lengthy.
   A scheme of argument, to be followed by the detail for the various parts of the scheme, will be given later.

   For the moment we will explore various theoretical consequences of Theorem (1) in the light of other established results concerned with:—

   - the (algebraic properties) of matrix inverses (Theorem (2), Theorem (3), Theorem (4), Theorem (5)),
   - the notions of linear combinations, linear dependence, and linear independence for column vectors and row vectors (Theorem (6), Theorem (7), Theorem (8)).

3. Recall the result below, labelled Theorem ($\sharp$) here, that has been already proved (and applied in other situations):

   Theorem ($\sharp$).

   Let A, B be $(p \times p)$-matrices. Suppose each of A, B is invertible. Then AB is invertible.

4. With the help of Theorem (1), we are going to prove the converse of Theorem ($\sharp$):

   Theorem (2). (Converse of Theorem ($\sharp$))

   Let A, B be $(p \times p)$-matrices.

   Suppose AB is invertible.

   Then each of A, B is invertible.

   Remark. Theorem (2) is not obvious, and when only rudimentary tools are allowed to be used, the argument for Theorem (2) will be non-trivial.

   However, with the help of Theorem (1), the argument for Theorem (2) is little more than simple matrix algebra.

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

   Let A, B be $(p \times p)$-matrices.

   Suppose AB is invertible. Denote the matrix inverse of AB by C.

   By definition, we have $(AB)C = I_p$ and $C(AB) = I_p$.

   We have $A(BC) = (AB)C = I_p$. Then A has a right inverse, namely BC. Hence, by Theorem (1), A is invertible.

   We also have $(CA)B = C(AB) = I_p$. Then B has a left inverse, namely, CA. Hence, by Theorem (1), B is invertible.

6. We may combine Theorem ($\sharp$) and Theorem (2) together to obtain Theorem (3). Further applying mathematical induction, we can deduce Theorem (4). A special case of Theorem (4) is Theorem (5).

   Theorem (3). (Invertibility of individual square matrices and their products.)

   Suppose A, B are $(p \times p)$-matrices.

   Then AB is invertible if and only if each of A, B is invertible.

   Theorem (4). (Corollary (1) to Theorem (3).)

   Suppose $A_1, A_2, \cdots, A_n$ are $(p \times p)$-matrices.

   Then $A_1 A_2 \cdots A_n$ is invertible if and only if each of $A_1, A_2, \cdots, A_n$ is invertible.

   Theorem (5). (Corollary (2) to Theorem (3).)

   Suppose A is a $(p \times p)$-matrix, and n is a positive integer.

   Then A^n is invertible if and only if A is invertible.

7. We recall the theoretical result below, labelled Theorem ($\natural_1$), about linear independence.

   Theorem ($\natural_1$).

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ are column vectors with p entries, and U is the $(p\times q)$-matrix given by $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_q\ ]$. (Here p, q are not assumed to be the same.)

   Then the statements (LI), (LI_0) are logically equivalent:—

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

   (Or equivalently: *For any numbers $\alpha_1, \alpha_2, \cdots, \alpha_q$, if $\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2+\cdots+\alpha_q\mathbf{u}_q = \mathbf{0}_p$ then $\alpha_1 = \alpha_2 = \cdots = \alpha_q = 0$.*

   Or further equivalently: *For any column vector $\mathbf{t}$ with q entries, if $U\mathbf{t} = \mathbf{0}_p$ then $\mathbf{t} = \mathbf{0}_q$.*)

   (LI_0) *The homogeneous system $\mathcal{LS}(U,\ \mathbf{0}_p)$ has no non-trivial solution.*

8. Combining Theorem (1) and Theorem ($\natural_1$), we obtain the re-formulation of invertibility in terms of linear independence.

   Theorem (6). (Re-formulation of invertibility in terms of linear independence.)

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

   (a) *A is invertible.*

   (d) *A has a left inverse.*

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

   (h) *For any column vector $\mathbf{t}$ with p entries, if $A\mathbf{t} = \mathbf{0}_p$ then $\mathbf{t} = \mathbf{0}_p$.*

   (j) *The columns of A are linearly independent.*

9. Now recall the theoretical result below, labelled Theorem ($\natural_2$), about linear combinations.

   Theorem ($\natural_2$).

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ are column vectors with p entries, and U is the $(p\times q)$-matrix given by $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_q\ ]$. (Here p, q are not assumed to be the same.)

   Suppose $\mathbf{v}$ is a column vector with p entries. Then the statements $(LC), (LC_0)$ are logically equivalent:—

   (LC) *$\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$.*

   (Or equivalently: *There exist some numbers $\alpha_1, \alpha_2, \cdots, \alpha_q$ such that $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_q\mathbf{u}_q$.*

   Or further equivalently: *There exists some column vector $\mathbf{t}$ with q entries such that $\mathbf{v} = U\mathbf{t}$.*)

   (LC_0) *The system $\mathcal{LS}(U,\ \mathbf{v})$ is consistent.*

10. Combining Theorem (1) and Theorem ($\natural_2$), we obtain the re-formulation of invertibility in terms of linear combinations.

Theorem (7). (Re-formulation of invertibility in terms of linear combinations.)

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

(a) *A is invertible.*

(e) *A has a right inverse.*

(g) *For any column vector $\mathbf{b}$ with p entries, the system $\mathcal{LS}(A, \mathbf{b})$ is consistent.*

(i) *For any column vector $\mathbf{b}$ with p entries, there exists some column vector $\mathbf{t}$ with p entries such that $A\mathbf{t} = \mathbf{b}$.*

(k) *Every column vector with p entries is a linear combination of the columns of A.*

11. We next recall the theoretical results below, labelled Theorem ($\flat$), about transpose and invertibility.

Theorem ($\flat$).

Suppose B is a $(p \times p)$-square matrix. Then the statements ($\flat$1), ($\flat$2) hold:

($\flat$1) *The equality $(B^t)^t = B$ holds.*

($\flat$2) *Suppose B is invertible. Then B^t is invertible, and the matrix inverse of B^t is given by $(B^t)^{-1} = (B^{-1})^t$.*

12. Combining Theorem (1), Theorem (6), Theorem (7) and Theorem ($\flat$), we obtain the result below, which gives a vastly expanded list of possible re-formulations (some of them highly non-obvious at first sight) for the notion of invertibility.

Theorem (8). (Various re-formulations for the notion of invertibiltiy.)

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

(a) *A is invertible.*

(a*) *A^t is invertible.*

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

(b*) *A^t is row-equivalent to I_p.*

(c) *A is a product of $(p \times p)$-row-operation matrices.*

(c*) *A^t is a product of $(p \times p)$-row-operation matrices.*

(d) *A has a left inverse.*

(d*) *A^t has a left inverse.*

(e) *A has a right inverse.*

(e*) *A^t has a right inverse.*

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

(f*) *The homogeneous system $\mathcal{LS}(A^t, \mathbf{0}_p)$ has no non-trivial solution.*

(g) *For any column vector $\mathbf{b}$ with p entries, the system $\mathcal{LS}(A, \mathbf{b})$ is consistent.*

(g*) *For any column vector $\mathbf{c}$ with p entries, the system $\mathcal{LS}(A^t, \mathbf{c})$ is consistent.*

(h) *For any column vector $\mathbf{t}$ with p entries, if $A\mathbf{t} = \mathbf{0}_p$ then $\mathbf{t} = \mathbf{0}_p$.*

(h*) *For any row vector $\mathbf{u}$ with p entries, if $\mathbf{u}A = {\mathbf{0}_p}^t$ then $\mathbf{u} = {\mathbf{0}_p}^t$.*

(i) *For any column vector $\mathbf{b}$ with p entries, there is some column vector $\mathbf{t}$ with p entries such that $A\mathbf{t} = \mathbf{b}$.*

(i*) *For any row vector $\mathbf{c}$ with p entries, there is some row vector $\mathbf{u}$ with p entries such that $\mathbf{u}A = \mathbf{c}$.*

(j) *The columns of A are linearly independent.*

(j*) *The rows of A are linearly independent.*

(k) *Every column vector with p entries is a linear combination of the columns of A.*

(k*) *Every row vector with p entries is a linear combination of the rows of A.*

Now suppose any one of the above holds (and hence all hold). Then the statements below hold:—

(α) *$[\ A \mid I_p\ ]$ is row-equivalent to $[\ I_p \mid A^{-1}\ ]$.*

(α^*) *$[\ A^t \mid I_p\ ]$ is row-equivalent to $[\ I_p \mid (A^t)^{-1}\ ]$.*

(β) *I_p is the only reduced row-echelon form which is row-equivalent to A.*

(β^*) *I_p is the only reduced row-echelon form which is row-equivalent to A^t.*

(γ) *For any column vector $\mathbf{b}$ with p entries, the system $\mathcal{LS}(A, \mathbf{b})$ has a unique solution, namely $A^{-1}\mathbf{b}$.*

(γ^*) *For any column vector $\mathbf{c}$ with p entries, the system $\mathcal{LS}(A^t, \mathbf{c})$ has a unique solution, namely $(A^t)^{-1}\mathbf{c}$.*

13. **Scheme of argument for Theorem (1).**

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

I. *Observations, on where we do not have to do anything.*

I.1. From what we have learnt about row operations and row-operation matrices, we know that the statement (b) and the statement (c) are logically equivalent:—

$$\underbrace{\textit{A is row-equivalent to } I_p}_{\text{Statement (b)}} \textit{ if and only if } \underbrace{\textit{A is a product of } (p \times p)\textit{-row-operation matrices}}_{\text{Statement (c)}}.$$

I.2. According to definition for the notion of invertibility, we know that the statement (a) implies the statement (d) and the statement (e):—

$$\textit{If } \underbrace{\textit{A is invertible}}_{\text{Statement (a)}}\textit{, then } \underbrace{\textit{A has a left inverse}}_{\text{Statement (d)}} \textit{ and } \underbrace{\textit{A has a right inverse}}_{\text{Statement (e)}}.$$

I.3. We have already known that the statement (a) implies the statement (γ):—

$$\textit{If } \underbrace{\textit{A is invertible}}_{\text{Statement (a)}} \textit{ then}$$

$$\underbrace{\textit{for any column vector } \mathbf{b} \textit{ with } p \textit{ entries, the system } \mathcal{LS}(A, \mathbf{b}) \textit{ has a unique solution, namely } A^{-1}\mathbf{b}}_{\text{Statement } (\gamma)}.$$

I.4. We have also known that the statement (d) implies the statement (f):—

$$\textit{If } \underbrace{\textit{A has a left inverse}}_{\text{Statement (d)}} \textit{ then } \underbrace{\textit{The homogeneous system } \mathcal{LS}(A, \mathbf{0}_p) \textit{ has no non-trivial solution.}}_{\text{Statement (f)}}$$

I.5. We have further known that the statement (e) implies the statement (g):—

$$\textit{If } \underbrace{\textit{A has a right inverse}}_{\text{Statement (e)}} \textit{ then}$$

$$\underbrace{\textit{for any column vector } \mathbf{b} \textit{ with } p \textit{ entries, the system } \mathcal{LS}(A, \mathbf{b}) \textit{ is consistent.}}_{\text{Statement (g)}}$$

II. *Observations, on where we need to do something.*

In the light of the above observations, we only have to prove three results, to be labelled Theorem (9), Theorem (10) and Theorem (11), whose purposes are explained below:—

II.1. In Theorem (9), we deduce that the statement (b) implies the statement (a) and the statement (α):—

$$\textit{If } \underbrace{\textit{A is row-equivalent to } I_p}_{\text{Statement (b)}}\textit{, then } \underbrace{\textit{A is invertible}}_{\text{Statement (a)}} \textit{ and } \underbrace{[\ A \mid I_p\ ] \textit{ is row-equivalent to } [\ I_p \mid A^{-1}\ ]}_{\text{Statement } (\alpha)}.$$

II.2. In Theorem (10), we deduce that that the statement (f) implies the statement (b):—

$$\textit{If } \underbrace{\textit{the homogeneous system } \mathcal{LS}(A, \mathbf{0}_p) \textit{ has no non-trivial solution}}_{\text{Statement (f)}} \textit{ then}$$

$$\underbrace{\textit{A is row-equivalent to } I_p}_{\text{Statement (b)}}.$$

II.3. By making a slight adapation to the argument for Theorem (10), we also deduce that the statement (f) implies the statement (β):—

$$\textit{If } \underbrace{\textit{the homogeneous system } \mathcal{LS}(A, \mathbf{0}_p) \textit{ has no non-trivial solution}}_{\text{Statement (f)}} \textit{ then}$$

$$\underbrace{I_p \textit{ is the only reduced row-echelon form which is row-equivalent to } A.}_{\text{Statement } (\beta)}$$

II.4. In Theorem (11), we deduce that the statement (g) implies the statement (b):—

$$\textit{If } \underbrace{\textit{for any column vector } \mathbf{b} \textit{ with } p \textit{ entries, the system } \mathcal{LS}(A, \mathbf{b}) \textit{ is consistent}}_{\text{Statement (g)}} \textit{ then}$$

$$\underbrace{\textit{A is row-equivalent to } I_p}_{\text{Statement (b)}}.$$

14. **Theorem (9).**

Let A be a $(p \times p)$-square matrix. Suppose A is row-equivalent to I_p.

Then the statements below hold:

(1) *A is invertible.*

(2) $[\ A \mid I_p\ ]$ *is row-equivalent to* $[\ I_p \mid A^{-1}\ ]$.

15. **Proof of Theorem (9).**

    Let A be a $(p \times p)$-square matrix. Suppose A is row-equivalent to I_p.

    Then A is a product of row-operation matrices, say, $G_1, G_2, \cdots, G_{k-1}, G_k$, giving some equality, say,

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

    Each of $G_1, G_2, \cdots, G_{k-1}, G_k$ is invertible. Then, since the equality $A = G_1 G_2 \cdots G_{k-1} G_k$ holds, A is also invertible.

    Moreover, $I_p = G_1 G_2 \cdots G_{k-1} G_k A^{-1}$.

    Hence we have

    $$[\ A \mid I_p\ ] = [\ AI_p \mid AA^{-1}\ ] = A[\ I_p \mid A^{-1}\ ] = G_1 G_2 \cdots G_{k-1} G_k [\ I_p \mid A^{-1}\ ]$$

    For each $j = 1, 2, \cdots, k, k-1$, denote by ρ_j the row-operation to which the row-operation matrix G_j is associated.

    Then we have the sequence of row operations joining $[\ I_p \mid A^{-1}\ ]$ to $[\ A \mid I_p\ ]$ below:

    $$[\ I_p \mid A^{-1}\ ] \xrightarrow{\rho_k} \xrightarrow{\rho_{k-1}} \xrightarrow{\rho_{k-2}} \cdots\cdots\cdots \xrightarrow{\rho_2} \xrightarrow{\rho_3} \xrightarrow{\rho_1} [\ A \mid I_p\ ]$$

    Therefore $[\ I_p \mid A^{-1}\ ]$ is row-equivalent to $[\ A \mid I_p\ ]$.

    Hence $[\ A \mid I_p\ ]$ is row-equivalent to $[\ I_p \mid A^{-1}\ ]$.

16. Recall the result below, denoted by Theorem (†) here, about systems of linear equations.

    Theorem (†).

    Let A, A' *be* $(p \times q)$*-matrices, and* $\mathbf{b}, \mathbf{b}'$ *be column vectors with* p *entries.*

    Suppose $A, \mathbf{b}$ *and* $A', \mathbf{b}'$ *are row-equivalent under the same sequence of row operations.*

    Suppose $\mathbf{t}$ *is a column vector with* q *entries.*

    Then $\mathbf{t}$ *is a solution of* $\mathcal{LS}(A,\ \mathbf{b})$ *if and only if* $\mathbf{t}$ *is a solution of* $\mathcal{LS}(A',\ \mathbf{b}')$.

17. With the help of Theorem (†), we establish Theorem (10).

    Theorem (10).

    Let A *be a* $(p \times p)$*-square matrix.*

    Suppose the homogenous system $\mathcal{LS}(A,\ \mathbf{0}_p)$ *has no non-trivial solution.*

    Then A *is row-equivalent to* I_p.

18. **Proof of Theorem (10).**

    Let A be a $(p \times p)$-square matrix.

    Suppose the homogenous system $\mathcal{LS}(A,\ \mathbf{0}_p)$ has no non-trivial solution.

    For such a matrix A, there is some reduced row-echelon form A' such that A' is row-equivalent to A.

    [We want to deduce that $A' = I_p$.]

    Note that $A, \mathbf{0}_p$ are row-equivalent to $A', \mathbf{0}_p$ under the same sequence of row operations.

    By Theorem (†), since $\mathcal{LS}(A,\ \mathbf{0}_p)$ has no non-trivial solution, $\mathcal{LS}(A',\ \mathbf{0}_p)$ also has no non-trivial solution.

    Then every column of A' is a pivot column.

    Therefore, for each $j = 1, 2, \cdots, p$, the (j,j)-th entry of A' is 1; all other entries of A' are 0.

    Since A' is a square matrix, we have $A' = I_p$.

    It follows that A is row-equivalent to I_p.

19. **Comment on the proof of Theorem (10).**

    By replacing the line

    > '*For such a matrix* A*, there is some reduced row-echelon form* A' *such that* A' *is row-equivalent to* A'

    with

'*Suppose A' a reduced row-echelon form which is row-equivalent to A*',

we obtain a passage which gives an argument for this result:—

> '*Let A be a $(p \times p)$-square matrix.*
> *Suppose the homogenous system $\mathcal{LS}(A,\ \mathbf{0}_p)$ has no non-trivial solution.*
> *Then I_p is the only reduced row-echelon form which is row-equivalent to A.*'

20. Recall the result below, denoted by Theorem (‡) here, about row operations and row-operation matrices.

    Theorem (‡).

    Let B, B' be $(p \times q)$-matrices.

    Suppose B is row-equivalent to B'.

    Then there are some $(p \times p)$-row-operation matrices $H_1, H_2, \cdots, H_k$ such that $B' = H_1 H_2 \cdots H_k B$.

21. With the help of Theorem (‡), we establish Theorem (11).

    Theorem (11).

    Let A be a $(p \times p)$-square matrix.

    Suppose that for any column vector $\mathbf{b}$ with p entries, the system $\mathcal{LS}(A,\ \mathbf{b})$ is consistent.

    Then A is row-equivalent to I_p.

22. **Proof of Theorem (11).**

    Let A be a $(p \times p)$-square matrix.

    Suppose that for any column vector $\mathbf{b}$ with p entries, the system $\mathcal{LS}(A,\ \mathbf{b})$ is consistent.

    - *Preparation.*
      For each $j = 1, 2, \cdots, p$, denote by $\mathbf{e}_j$ the column vector with j-th entry being 1 and with all other entries being 0.
      Then by assumption, for each $j = 1, 2, \cdots, p$, the system $\mathcal{LS}(A,\ \mathbf{e}_j)$ is consistent, say, with solution $\mathbf{u}_j$.
      Define the $(p \times p)$-matrix U by $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_p\ ]$.
      We have the chain of equalities $AU = A[\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_p\ ] = [\ A\mathbf{u}_1 \mid A\mathbf{u}_2 \mid \cdots \mid A\mathbf{u}_p\ ] = [\ \mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_p\ ] = I_p$.
    - *More preparation.*
      There is some reduced row-echelon form A', of rank, say, r, such that A' is row-equivalent to A.
      [We want to deduce that $r = p$ and $A' = I_p$.]
      By Theorem (‡), since A is row-equivalent to A', there are some $(p \times p)$-row-operation matrices $H_1, H_2, \cdots, H_k$ such that $A' = H_1 H_2 \cdots H_k A$.
      Write $H = H_1 H_2 \cdots H_k$. Then $A' = HA$.
      Each of $H_1, H_2, \cdots, H_k$ is invertible. Then H is invertible.
      Note that $A'U = (HA)U = H(AU) = HI_p = H$.
      Since H is invertible, we have $A'UH^{-1} = HH^{-1} = I_p$.

    We claim that every row of A' is a non-zero row.

    This is justified, with the proof-by-contradiction argument:

    - Suppose that not every row of A' were a non-zero row.
      Then the bottom row of A' is a row of 0's.
      By the definition of matrix multiplication, the bottom row of $A'UH^{-1}$ is also a row of 0's.
      But $A'UH^{-1} = I_p$. So the last entry of $A'UH^{-1}$ is 1. This is impossible.

    So we have shown that every row of A' is a non-zero row. Then $r = p$.

    Therefore all p columns of A' are pivot columns. It follows that $A' = I_p$.

    Hence A is row-equivalent to I_p.