4.6.1 Appendix: Proofs of basic theoretical results concerned with column space, row space, and minimal spanning set.

0. The material in this appendix is supplementary.

1. We have introduced Theorem (1), and we are going to give a proof for this result.

   Theorem (1). ('Minimal spanning set' and dimension for a span of several column vectors.)

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q \in \mathbb{R}^n$, *and* $\mathcal{W} = \text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$.

   Define $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_q\ ]$, *and denote the rank of* U *by* r.

   Then the statements below hold:—

   (a) $\dim(\mathcal{W}) = r$.

   (b) *From now on suppose* $r \geq 1$.

   Suppose $U^\sharp$ *is a row-echelon form which is row-equivalent to* U, *and suppose the pivot columns of* $U^\sharp$, *from left to right, are the* d_1*-th,* d_2*-th, ...,* d_r*-th columns of* $U^\sharp$.

   Then $\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}$ *constitute a basis for* $\mathcal{W}$ *over the reals.*

   (c) *Suppose* U' *is the reduced row-echelon form which is row-equivalent to* U.

   For each $j = 1, 2, \cdots, q$, *denote the* j*-th column of* U' *by* $\mathbf{u}'_j$, *and denote the first* r *entries, from the top downwards, by* $\mathbf{u}'_j$ *by* $\alpha_{1j}, \alpha_{2j}, \cdots, \alpha_{rj}$.

   Suppose the k*-th column of* U' *is a free column in* U'.

   Then $\mathbf{u}_k = \alpha_{1k}\mathbf{u}_{d_1} + \alpha_{2k}\mathbf{u}_{d_2} + \cdots + \alpha_{rk}\mathbf{u}_{d_r}$.

2. In the argument for Theorem (1), we will have to apply a few previously established results. They are stated as reference here:—

   Theorem (⋆).

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q \in \mathbb{R}^n$.

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ *are linearly independent over the reals.*

   Then:—

   (a) *the inequality* $q \leq n$ *holds, and*

   (b) $\text{Span}(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$ *is a* q*-dimensional subspace of* $\mathbb{R}^n$ *over the reals, with a basis over the reals constituted by* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$.

   Theorem (⋆⋆).

   Suppose $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r, \mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_s \in \mathbb{R}^n$. *Then the statements below are logically equivalent:—*

   (1) *Each of* $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_s$ *is a linear combination of* $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r$ *over the reals.*

   (2) $\text{Span}(\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r, \mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_s\}) = \text{Span}(\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r\})$.

   Theorem ($\sharp_1$). ('Preservation' of linear combinations by left-multiplication by invertible matrices.)

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q, \mathbf{v}$ *be column vectors with* m *entries.*

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

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

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

   Theorem ($\sharp_2$). ('Preservation' of linear independence by left-multiplication by invertible matrices.)

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ *be column vectors with* m *entries.*

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

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

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

   We are now ready to give a proof for Theorem (1).

3. **Proof of Theorem (1).**

Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q \in \mathbb{R}^n$, and $\mathcal{W} = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$.

Define $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_q\ ]$, and denote the rank of U by r.

If $r = 0$, then $U = 0$ and $\mathcal{W} = \{\mathbf{0}_n\}$ and $\dim(\mathcal{W}) = 0$.

From now on we assume $r \geq 1$.

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

For each $j = 1, 2, \cdots, q$, denote the j-th column of U' by $\mathbf{u}'_j$, and denote the first r entries, from the top downwards, by $\mathbf{u}'_j$ by $\alpha_{1j}, \alpha_{2j}, \cdots, \alpha_{rj}$.

(a) Since U is row-equivalent to U', there exists some invertible $(n \times n)$-square matrix G such that $U' = GU$.
For the same G, we have $\mathbf{u}'_j = G\mathbf{u}_j$ for each $j = 1, 2, \cdots, q$.

(b) By assumption, the rank of U' is r.
Suppose the pivot columns in U', from left to right, are the d_1-th, d_2-th, ..., d_r-th columns.
We are going to verify that $\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}$ constitute a basis for $\mathcal{W}$. It will then follow that $\dim(\mathcal{W}) = r$ as well.

   i. For each $\ell = 1, 2, \cdots, r$, we have $\mathbf{u}'_{d_\ell} = \mathbf{e}_\ell^{(n)}$.
   Then $\mathbf{u}'_{d_1}, \mathbf{u}'_{d_2}, \cdots, \mathbf{u}'_{d_r}$ are linearly independent over the reals.
   Therefore, by Theorem ($\sharp_2$), the column vectors $\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}$ are linearly independent over the reals.

   ii. For each $j = 1, 2, \cdots, q$, all entries in $\mathbf{u}'_j$ below its r-th entry are 0.
   Then $\mathbf{u}'_j = \alpha_{1j}\mathbf{e}_1^{(n)} + \alpha_{2j}\mathbf{e}_2^{(n)} + \cdots + \alpha_{rj}\mathbf{e}_r^{(n)} = \alpha_{1j}\mathbf{u}'_{d_1}\alpha_{2j}\mathbf{u}'_{d_2} + \cdots + \alpha_{rj}\mathbf{u}'_{d_r}$.
   Therefore $\mathbf{u}'_j$ is a linear combination of $\mathbf{u}'_{d_1}, \mathbf{u}'_{d_2}, \cdots, \mathbf{u}'_{d_r}$ over the reals, with a linear relation $\mathbf{u}'_j = \alpha_{1j}\mathbf{u}'_{d_1} + \alpha_{2j}\mathbf{u}'_{d_2} + \cdots + \alpha_{rj}\mathbf{u}'_{d_r}$.
   Hence, by Theorem ($\sharp_1$), the column vector $\mathbf{u}_j$ is a linear combination of $\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}$ over the reals, with a linear relation $\mathbf{u}_j = \alpha_{1j}\mathbf{u}_{d_1} + \alpha_{2j}\mathbf{u}_{d_2} + \cdots + \alpha_{rj}\mathbf{u}_{d_r}$.

   iii. Whenever k is amongst $1, 2, \cdots, q$ but not amongst $d_1, d_2, \cdots, d_r$, the column vector $\mathbf{u}_k$ is a linear combination of $\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}$.
   Then by Theorem ($\star\star$), we have the equalities

$$\mathcal{W} = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\}) = \mathsf{Span}\ (\{\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}\})$$

   Recall that $\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}$ are linearly independent over the reals.
   It follows from Theorem ($\star$) that $\mathcal{W}$ is an r-dimensional subspace of $\mathbb{R}^n$ over the reals, with a basis given by $\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}$.

(c) Suppose the k-th column of U' is a free column in U'. Then k is amongst $1, 2, \cdots, q$ but not amongst $d_1, d_2, \cdots, d_r$. As argued, $\mathbf{u}_k$ is a linear combination of $\mathbf{u}_{d_1}, \mathbf{u}_{d_2}, \cdots, \mathbf{u}_{d_r}$ over the reals, with a linear relation $\mathbf{u}_k = \alpha_{1k}\mathbf{u}_{d_1} + \alpha_{2k}\mathbf{u}_{d_2} + \cdots + \alpha_{rk}\mathbf{u}_{d_r}$.

4. Recall the definition for the notion of column space and row space, which are formulated in terms of span. We are going to prove a few theoretical results concerned with the notions of column space and row space.

Definition. (Column space and row space for matrices.)

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

(1) *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* $\mathsf{Span}\ (\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_q\})$. *It is denoted by $\mathcal{C}(A)$.*

(2) *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* $\mathsf{Span}\ (\{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_p\})$. *It is denoted by $\mathcal{R}(A)$.*

5. We take for granted Lemma (3), whose proof is left as an easy exercise.

Lemma (3).

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

(a) $\mathcal{C}(A^t) = \mathcal{R}(A)$, *and* $\mathcal{C}(A) = \mathcal{R}(A^t)$.

(b) i. $\mathcal{C}(A) = \left\{ \mathbf{y} \in \mathbb{R}^p \;\middle|\; \text{There exist some } \mathbf{t} \in \mathbb{R}^q \text{ such that } \mathbf{y} = A\mathbf{t}. \right\}$.

ii. $\mathcal{R}(A) = \left\{ \mathbf{x} \in \mathbb{R}^q \;\middle|\; \text{There exist some } \mathbf{s} \in \mathbb{R}^p \text{ such that } \mathbf{x}^t = \mathbf{s}^t A. \right\}$.

6. We have introduced a pair of results concerned with column spaces and row spaces. We are going to prove these results. They are exercises on the definition for the notion of set equality and that of invertibility for square matrices, together with some basic matrix algebra.

Theorem (4).

Let A, B be matrices with n rows and with real entries.

Suppose H is an invertible $(n \times n)$-square matrix with real entries. Then the statements below are logically equivalent:—

(1) $\mathcal{C}(A) = \mathcal{C}(B)$.

(2) $\mathcal{C}(HA) = \mathcal{C}(HB)$.

Theorem (5).

Let B be a matrix with n columns and with real entries, and G be an $(n \times n)$-square matrix with real entries.

Suppose G in invertible. Then $\mathcal{C}(B) = \mathcal{C}(BG)$.

7. **Proof of Theorem (4).**

Let A, B be matrices with n rows and with real entries. Suppose A has p columns, and B has q columns.

Suppose H is an invertible $(n \times n)$-square matrix.

(a) Suppose the statement (1) holds: $\mathcal{C}(A) = \mathcal{C}(B)$.
[We want to deduce the statement (2): $\mathcal{C}(HA) = \mathcal{C}(HB)$.]

i. [We want to verify the statement (†): *For any $\mathbf{y} \in \mathbb{R}^n$, if $\mathbf{y} \in \mathcal{C}(HA)$ then $\mathbf{y} \in \mathcal{C}(HB)$.*]
Pick any $\mathbf{y} \in \mathbb{R}^n$. Suppose $\mathbf{y} \in \mathcal{C}(HA)$.
[We want to deduce '$\mathbf{y} \in \mathcal{C}(HB)$'.
This amounts to deducing: '*there exists some $\mathbf{t} \in \mathbb{R}^q$ such that $\mathbf{y} = (HB)\mathbf{t}$.*'
Now we ask: How to name such an appropriate $\mathbf{t}$? How do the assumptions '$\mathcal{C}(A) = \mathcal{C}(B)$', '$\mathbf{y} \in \mathcal{C}(HA)$' help?]
By the definition of column space, there exists some $\mathbf{u} \in \mathbb{R}^p$ such that $\mathbf{y} = (HA)\mathbf{u}$.
For the same $\mathbf{u}$, we have $\mathbf{y} = H(A\mathbf{u})$.
Also by the definition of column space, $A\mathbf{u} \in \mathcal{C}(A)$.
Since $\mathcal{C}(A) = \mathcal{C}(B)$, we have $A\mathbf{u} \in \mathcal{C}(B)$.
Then by the definition of column space, there exists some $\mathbf{t} \in \mathbb{R}^q$ such that $A\mathbf{u} = B\mathbf{t}$.
Then, for the same $\mathbf{y}, \mathbf{u}, \mathbf{t}$, we have $\mathbf{y} = (HA)\mathbf{u} = H(A\mathbf{u}) = H(B\mathbf{t}) = (HB)\mathbf{t}$.
Therefore, by the definition of column space, $\mathbf{y} \in \mathcal{C}(HB)$.

ii. [We want to verify the statement (‡): *For any $\mathbf{z} \in \mathbb{R}^n$, if $\mathbf{z} \in \mathcal{C}(HB)$ then $\mathbf{z} \in \mathcal{C}(HA)$.*]
Pick any $\mathbf{z} \in \mathbb{R}^n$. Suppose $\mathbf{z} \in \mathcal{C}(HB)$.
By the definition of column space, there exists some $\mathbf{s} \in \mathbb{R}^q$ such that $\mathbf{z} = (HB)\mathbf{s}$.
For the same $\mathbf{s}$, we have $\mathbf{z} = H(B\mathbf{s})$.
Also by the definition of column space, $B\mathbf{s} \in \mathcal{C}(B)$.
Since $\mathcal{C}(A) = \mathcal{C}(B)$, we have $B\mathbf{s} \in \mathcal{C}(A)$.
Then by the definition of column space, there exists some $\mathbf{v} \in \mathbb{R}^p$ such that $B\mathbf{s} = A\mathbf{v}$.
Then, for the same $\mathbf{z}, \mathbf{s}, \mathbf{v}$, we have $\mathbf{z} = (HB)\mathbf{s} = H(B\mathbf{s}) = H(A\mathbf{v}) = (HA)\mathbf{v}$.
Therefore, by the definition of column space, $\mathbf{z} \in \mathcal{C}(HA)$.

It follows that $\mathcal{C}(HA) = \mathcal{C}(HB)$.

(b) Suppose the statement (2) holds: $\mathcal{C}(HA) = \mathcal{C}(HB)$.

Since H is invertible, the matrix inverse H^{-1} is well-defined as an invertible $(n \times n)$-square matrix.

From the reasoning above (with the roles of 'A, B, H' taken by 'HA, HB, H^{-1}' respectively), we deduce the equality $\mathcal{C}(H^{-1}(HA)) = \mathcal{C}(H^{-1}(HB))$ from the equality $\mathcal{C}(HA) = \mathcal{C}(HB)$.

It then follows (from the equalities '$A = H^{-1}(HA)$', '$B = H^{-1}(HB)$') that

$$\mathcal{C}(A) = \mathcal{C}\big(H^{-1}(HA)\big) = \mathcal{C}\big(H^{-1}(HB)\big) = \mathcal{C}(B)\,.$$

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

   Let B be a $(q \times n)$-matrix with real entries, and G is an $(n \times n)$-square matrix with real entries.

   Suppose G is invertible.

   (a) [We want to verify the statement (†): *For any* $\mathbf{y} \in \mathbb{R}^q$, *if* $\mathbf{y} \in \mathcal{C}(BG)$ *then* $\mathbf{y} \in \mathcal{C}(B)$.]

   Pick any $\mathbf{y} \in \mathbb{R}^q$. Suppose $\mathbf{y} \in \mathcal{C}(BG)$.

   [We want to deduce '$\mathbf{y} \in \mathcal{C}(B)$'.
   This amounts to deducing: '*there exists some* $\mathbf{t} \in \mathbb{R}^n$ *such that* $\mathbf{y} = B\mathbf{t}$.'
   Now we ask: How to name such an appropriate $\mathbf{t}$? How to make use of the assumptions '$\mathbf{y} \in \mathcal{C}(BG)$'?]

   By the definition of column space, there exists some $\mathbf{s} \in \mathbb{R}^n$ such that $\mathbf{y} = (BG)\mathbf{s}$.

   Define $\mathbf{t} = G\mathbf{s}$. By definition, $\mathbf{t} \in \mathbb{R}^n$.

   For the same $\mathbf{y}, \mathbf{s}, \mathbf{t}$, we have $\mathbf{y} = (BG)\mathbf{s} = B(G\mathbf{s}) = B\mathbf{t}$.

   Then, by the definition of column space, we have $\mathbf{y} \in \mathcal{C}(B)$.

   (b) [We want to verify the statement (‡): *For any* $\mathbf{z} \in \mathbb{R}^n$, *if* $\mathbf{z} \in \mathcal{C}(B)$ *then* $\mathbf{z} \in \mathcal{C}(BG)$.]

   Pick any $\mathbf{z} \in \mathbb{R}^q$. Suppose $\mathbf{z} \in \mathcal{C}(G)$.

   [We want to deduce '$\mathbf{z} \in \mathcal{C}(BG)$'.
   This amounts to deducing: '*there exists some* $\mathbf{s} \in \mathbb{R}^n$ *such that* $\mathbf{z} = (BG)\mathbf{s}$.'
   Now we ask: How to name such an appropriate $\mathbf{s}$? How to make use of the assumptions '$\mathbf{z} \in \mathcal{C}(BG)$'?]

   By definition of column space, there exists some $\mathbf{t} \in \mathbb{R}^n$ such that $\mathbf{y} = B\mathbf{t}$.

   Since G is invertible, its matrix inverse G^{-1} is well-defined as an $(n \times n)$-square matrix.

   Define $\mathbf{s} = G^{-1}\mathbf{t}$. By definition $\mathbf{s} \in \mathbb{R}^n$.

   For the same $\mathbf{z}, \mathbf{t}, \mathbf{s}$, we have $\mathbf{z} = B\mathbf{t} = B(I_n\mathbf{t}) = B[(GG^{-1})\mathbf{t}] = B[G(G^{-1}\mathbf{t})] = B(G\mathbf{s}) = (BG)\mathbf{s}$.

   Then, by the definition of column space, we have $\mathbf{z} \in \mathcal{C}(BG)$.

   It follows that $\mathcal{C}(B) = \mathcal{C}(BG)$.

9. We have explained that an immediate consequence of Theorem (5) is the pair of results below, namely, Theorem (6) and Theorem (7).

   Theorem (6). (Corollary to Theorem (5).)

   Let A be a matrix with n rows and with real entries, and H be an $(n \times n)$-square matrix with real entries.

   Suppose H in invertible. Then $\mathcal{R}(A) = \mathcal{R}(HA)$.

   Theorem (7). (Corollary to Theorem (6).)

   Let $A, \widetilde{A}$ be matrices with real entries of the same size.

   Suppose A is row-equivalent to $\widetilde{A}$.

   Then $\mathcal{R}(A) = \mathcal{R}\left(\widetilde{A}\right)$.

10. We have introduced the important theoretical result, namely, Theorem (2), which says the respective dimensions of the column space and the row space of an arbitrary matrix are the same as the rank of that matrix. We are going to givea proof for Theorem (2).

    Theorem (2). (Equality amongst rank, 'column rank' and 'row rank' for an arbitrarily given matrix.)

    Let A be an $(n \times p)$-matrix with real entries.

    Suppose A is of rank r.

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

11. The proof of Theorem (2) relies on Theorem (1), Theorem (7), and also the technical (but easy) result below about linear independence and bases. The proof of Lemma (♮) is left as an exercise on the definitions for the notions of linear independence and basis.

    Lemma (♮).

    Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ be non-zero column vectors with m real entries.

    Suppose that for each $j = 1, 2, \cdots, k$, the first non-zero entry in $\mathbf{v}_{j+1}$ is strictly below the first non-zero entry in $\mathbf{v}_j$.

    Then $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly independent over the reals, and constitute a basis for $\mathrm{Span}(\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\})$.

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

Let A be an $(n \times p)$-matrix with real entries.

Suppose A is of rank r.

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

By definition, there are:—

- r non-zero rows in A',
- r leading ones in A', and
- r pivot columns in A'.

(a) Label the columns of A, from left to right, by $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p$.
Suppose the pivot columns of A' are the d_1-th, d_2-th, ..., d_r-th columns of A'.
By Lemma (3), we have $\mathcal{C}(A) = \text{Span}(\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p\})$.
By Theorem (1), $\mathbf{a}_{d_1}, \mathbf{a}_{d_2}, \cdots, \mathbf{a}_{d_r}$ constitute a basis for $\text{Span}(\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p\})$ over the reals.
Then $\dim(\mathcal{C}(A)) = \dim(\text{Span}(\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p\})) = r$.

(b) We now proceed to verify that $\dim(\mathcal{R}(A)) = r$.
By Theorem (7), since A is row-equivalent to A', we have $\mathcal{R}(A) = \mathcal{R}(A')$.
We now display a basis with r column vectors for $\mathcal{R}(A')$ over the reals:—

- Write $B = (A')^t$. The only non-zero columns of B are the first r columns of B. Label them, from left to right, by $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r$. They are the respective transposes of the 1-st, 2-nd, ..., r-th rows of A'.
  For each $j = 1, 2, \cdots, r$, the first non-zero entry of the $(j+1)$-th row of A' is strictly to the right of the first non-zero entry of the j-th row of A'.
  Then the first non-zero entry of $\mathbf{b}_{j+1}$ is strictly below the first non-zero entry of $\mathbf{b}_j$.
  Then by Lemma ($\natural$), $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r$ constitute a basis for $\text{Span}(\{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r\})$ over the reals.
  Recall that the columns of B other than the first r columns are all given by $\mathbf{0}_p$ (which is trivially a linear combination of $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r$).
  Then by the definition of row space and column space, we have the equalities

$$\mathcal{R}(A') = \mathcal{C}(B) = \text{Span}(\{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r, \mathbf{0}_p, \mathbf{0}_p, \cdots, \mathbf{0}_p\}) = \text{Span}(\{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r\}).$$

  (The last equality is due to Theorem ($\star\star$).)
  Therefore $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_r$ constitute a basis for $\mathcal{R}(A')$.

Now we have $\dim(\mathcal{R}(A)) = \dim(\mathcal{R}(A')) = r$.