Lemma 3.3.1. Let \( \{a^i\}_{i=1}^k = \{[a^i_1, a^i_2, \ldots, a^i_n]\}_{i=1}^k \) be a set of vectors in \( \mathbb{R}^n \) and let \( \pi \in S_n \). Then the set of vectors \( \{a^i\}_{i=1}^k \) is linearly independent if and only if the set of vectors \( \{[a^i_{\pi(1)}, a^i_{\pi(2)}, \ldots, a^i_{\pi(n)}]\}_{i=1}^k \) is linearly independent. That is, permuting all the entries in a set of vectors by the same permutation preserves the linear dependence/independence of the set.

**Proof.** Set \( \{a^i\}_{i=1}^k \) is linearly independent if and only if \( \sum_{i=1}^k s_i a^i = 0 \) for scalars \( s_1, s_2, \ldots, s_k \) then \( s_1 = s_2 = \cdots = s_k = 0 \). Now \( \sum_{i=1}^k s_i a^i = 0 \) implies that \( \sum_{i=1}^k s_i a^i_j = 0 \) for \( j = 1, 2, \ldots, n \). So this system of \( n \) linear equations (in \( k \) unknowns \( s_i \) for \( i = 1, 2, \ldots, k \)) has only one solution if and only if the system of \( n \) linear equations in \( k \) unknowns \( \sum_{i=1}^k s_i a^i_{\pi(j)} = 0 \) for \( j = 1, 2, \ldots, n \) has only one solution, namely \( s_1 = s_2 = \cdots = s_k = 0 \). That is, if and only if the vector equation \( \sum_{i=1}^k s_i b^i = 0 \), where \( b^i = [a^i_{\pi(1)}, a^i_{\pi(2)}, \ldots, a^i_{\pi(n)}] \) for \( i = 1, 2, \ldots, k \), has only one solution, namely \( s_1 = s_2 = \cdots = s_k = 0 \).

Lemma 3.3.1 (continued)

Theorem 3.3.2

**Theorem 3.3.2.** Let \( A \) be an \( n \times m \) matrix. Then the row rank of \( A \) equals the column rank of \( A \). This common quantity is called the rank of \( A \).

**Proof.** Let the row rank of \( A \) be \( p \) and let the column rank of \( A \) be \( q \). Rearrange the rows of \( A \) to form matrix \( B \) so that the first \( p \) rows of matrix \( B \) are linearly independent (so \( B = PA \) where \( P \) is some permutation matrix). Since \( A \) and \( B \) have the same rows, they have equal row rank. By Lemma 3.3.1, the column rank of \( A \) equals the column rank of \( B \) (by interchanging row \( i \) and \( j \) of \( A \), we are interchanging all of the \( i \)th entries with the \( j \)th entries in the column vectors of \( A \)). So we can partition \( B \) as \( B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \) where the \( p \) rows of \( B_1 \) are linearly independent and the \( n - p \) rows of \( B_2 \) are (each) linear combinations of the rows of \( B_1 \). So with the rows of \( B_1 \) as \( r_1, r_2, \ldots, r_p \) and the rows of \( B_2 \) as \( r_{p+1}, r_{p+2}, \ldots, r_n \), we have scalars \( s_{\ell i} \) where \( r_\ell = \sum_{i=1}^p s_{\ell i} r_i \) for \( \ell = p + 1, p + 2, \ldots, n \).
Theorem 3.3.2 (continued)

**Proof (continued).** Then with $S$ the $(n - p) \times p$ matrix with entries $s_{ij}$, $S = [s_{ij}]$, we have $B_2 = SB_1$. So $B = \begin{bmatrix} B_1 \\ SB_1 \end{bmatrix}$. We claim now that the column rank of $B$ is the same as the column rank of $B_1$.

With $s = [s_1, s_2, \ldots, s_m]^T$ as a vector of $m$ scalars, we have $Bs = 0$ if and only if $\begin{bmatrix} B_1 \\ SB_1 \end{bmatrix} s = \begin{bmatrix} B_1 s \\ SB_1 s \end{bmatrix} = 0$ if and only if $B_1 s = 0$. That is, a linear combination of the columns of $B$ is 0 if and only if the corresponding linear combination of the columns of $B_1$ is 0. So the column rank of $B$ is the same as the column rank of $B_1$, and so both are the same as the column rank of $A$ (namely, $q$). Since the columns of $B_1$ are vectors in $\mathbb{R}^p$ then $q \leq p$.

Similarly, we can rearrange the columns of $A$ and partition the resulting matrix to show that $p \leq q$. Therefore the row rank, $p$, of matrix $A$ equals the column rank, $q$, of matrix $A$. \qed

---

Theorem 3.3.3

If $P$ and $Q$ are products of elementary matrices then $\text{rank}(PAQ) = \text{rank}(A)$.

**Proof.** We show the result holds for $P$ a single elementary matrix. The result for $Q$ a single elementary matrix follows similarly and the general result then follows by induction. Let $P = E_{pq}$ where $I_n \xrightarrow{R_p \mapsto R_q} E_{pq}$. Then $E_{pq}A$ has the same rows as $A$ and so $\text{rank}(E_{pq}A) = \text{rank}(A)$. Let $P = E_{sp}$ where $l_n \xrightarrow{R_p \mapsto R_q} E_{sp}$ where $s \neq 0$. Then with $r_1, r_2, \ldots, r_n$ as the rows of $A$, we have that $r_1, r_2, \ldots, r_{p-1}, s_{p}, r_{p+1}, \ldots, r_n$ are the rows of $E_{sp}A$. Now

$$\sum_{i=1}^{n} s_i r_i = \sum_{i=1}^{p-1} s_i r_i + (s_p/s) s_p r_p + \sum_{i=p+1}^{n} s_i r_i$$

for any scalars $s_1, s_2, \ldots, s_n$. So $r_1, r_2, \ldots, r_n$ and $r_1, r_2, \ldots, r_{p-1}, s_{p}, r_{p+1}, \ldots, r_n$ satisfy precisely the same dependence/independence relations. Therefore $\text{rank}(E_{sp}A) = \text{rank}(A)$.

---

Theorem 3.3.3 (continued)

**Theorem 3.3.3.** If $P$ and $Q$ are products of elementary matrices then $\text{rank}(PAQ) = \text{rank}(A)$.

**Proof (continued).** Let $P = E_{pq}$ where $I_n \xrightarrow{R_p \mapsto R_q} E_{pq}$. Then for $r_1, r_2, \ldots, r_n$ the rows of $A$, we have that

$$r_1, r_2, \ldots, r_{p-1}, r_{p} + s_{q}, r_{p+1}, \ldots, r_n$$

are the rows of $E_{pq}A$. Now

$$\sum_{i=1}^{n} s_i r_i = \sum_{i=1}^{p-1} s_i r_i + (s_{p} + s) r_{p} + \sum_{i=p+1}^{n} s_i r_i$$

for any scalars $s_1, s_2, \ldots, s_n$. So $r_1, r_2, \ldots, r_n$ and $r_1, r_2, \ldots, r_{p-1}, r_{p} + s_{q}, r_{p+1}, \ldots, r_n$ satisfy precisely the same dependence/independence relations. Therefore $\text{rank}(E_{pq}A) = \text{rank}(A)$.

---

Theorem 3.3.4

**Theorem 3.3.4.** Let $A$ be a matrix partitioned as $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$. Then

(i) $\text{rank}(A_{ij}) \leq \text{rank}(A)$ for $i, j \in \{1, 2\}$.

(ii) $\text{rank}(A) \leq \text{rank}([A_{11} | A_{12}]) + \text{rank}([A_{21} | A_{22}])$.

(iii) $\text{rank}(A) \leq \text{rank} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right)$.

(iv) If $\mathcal{V}([A_{11} | A_{12}]^T) \perp \mathcal{V}([A_{21} | A_{22}]^T)$ then $\text{rank}(A) = \text{rank}([A_{11} | A_{12}]) + \text{rank}([A_{21} | A_{22}])$ and if

$$\mathcal{V} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) \perp \mathcal{V} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right)$$

then

$$\text{rank}(A) = \text{rank} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right).$$
Theorem 3.3.4 (continued 1)

(i) \( \text{rank}(A_{ij}) \leq \text{rank}(A) \) for \( i,j \in \{1,2\} \).

Proof. (i) Since the set of rows of \( [A_{11} | A_{12}] \) is a subset of the set of rows of \( A \), then by Exercise 2.1.G(i), \( \text{rank}([A_{11} | A_{12}]) \leq \text{rank}(A) \). Similarly, the set of columns of \( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \) is a subset of the set of columns of \( A \) and so

\[
\text{rank} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) \leq \text{rank}(A). \]

Also, \( \text{rank}([A_{21} | A_{22}]) \leq \text{rank}(A) \) and

\[
\text{rank} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right) \leq \text{rank}(A). \]

Next, the set of columns of \( A_{11} \) is a subset of the set of columns of \( [A_{11} | A_{12}] \) and so \( \text{rank}(A_{11}) \leq \text{rank}([A_{11} | A_{12}]) \) (and similarly \( \text{rank}(A_{12}) \leq \text{rank}([A_{11} | A_{12}]) \)). Therefore

\[
\text{rank}(A_{11}) \leq \text{rank}([A_{11} | A_{12}]) \leq \text{rank}(A) \text{ and } \text{rank}(A_{12}) \leq \text{rank}([A_{11} | A_{12}]) \leq \text{rank}(A). \]

Similarly, \( \text{rank}(A_{21}) \leq \text{rank}([A_{21} | A_{22}]) \leq \text{rank}(A) \) and

\[
\text{rank}(A_{22}) \leq \text{rank}([A_{21} | A_{22}]) \leq \text{rank}(A). \]

Theorem 3.3.4 (continued 2)

(ii) \( \text{rank}(A) \leq \text{rank}([A_{11} | A_{12}]) + \text{rank}([A_{21} | A_{22}]) \).

(iii) \( \text{rank}(A) \leq \text{rank} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right) \).

Proof (continued). (ii) Let \( R \) be the set of rows of \( A \), \( R_1 \) the set of rows of \( [A_{11} | A_{12}] \), and \( R_2 \) the set of rows of \( [A_{21} | A_{22}] \). Then \( R = R_1 \cup R_2 \) and by Exercise 2.1.G(ii), \( \text{dim}(\text{span}(R)) \leq \text{dim}(\text{span}(R_1)) + \text{dim}(\text{span}(R_2)) \).

That is, \( \text{rank}(A) \leq \text{rank}([A_{11} | A_{12}]) + \text{rank}([A_{21} | A_{22}]) \).

(iii) Let \( C \) be the set of columns of \( A \). \( C_1 \) be the set of columns of \( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \), and \( C_2 \) be the set of columns of \( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \). Then \( C = C_1 \cup C_2 \) and by Exercise 2.1.G(ii),

\[
\text{dim}(\text{span}(C)) \leq \text{dim}(\text{span}(C_1)) + \text{dim}(\text{span}(C_2)) \text{.} \]

That is,

\[
\text{rank}(A) \leq \text{rank} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right) \text{.} \]

Theorem 3.3.4 (continued 3)

(iv) If \( \mathcal{V}(\begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix}^T) \perp \mathcal{V}(\begin{bmatrix} A_{21} \\ A_{22} \end{bmatrix}^T) \) then

\[
\text{rank}(A) = \text{rank}(\begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix}) + \text{rank}(\begin{bmatrix} A_{21} \\ A_{22} \end{bmatrix}) \]

and if \( \mathcal{V} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) \perp \mathcal{V} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right) \) then

\[
\text{rank}(A) = \text{rank} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right) \text{.} \]

Proof (continued). (iv) Let \( R \) be the set of rows of \( A \), \( R_1 \) the set of rows of \( [A_{11} | A_{12}] \), and \( R_2 \) the set of rows of \( [A_{21} | A_{22}] \). Then \( \mathcal{V}(\begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix}^T) \) is the column space of \( [A_{11} | A_{12}] \) and \( \mathcal{V}(\begin{bmatrix} A_{21} \\ A_{22} \end{bmatrix}^T) \) is the column space of \( [A_{21} | A_{22}] \). So the column space of \( A \) is \( \mathcal{V}(\begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix}^T) \perp \mathcal{V}(\begin{bmatrix} A_{21} \\ A_{22} \end{bmatrix}^T) \) (see page 13 of the text). Since \( \mathcal{V}(\begin{bmatrix} A_{21} \\ A_{22} \end{bmatrix}^T) \perp \mathcal{V}(\begin{bmatrix} A_{21} | A_{22} \end{bmatrix}^T) \) by hypothesis, then the column space of \( A \) is \( \mathcal{V}(\begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix}^T) \perp \mathcal{V}(\begin{bmatrix} A_{21} | A_{22} \end{bmatrix}^T) \).

By Exercise 2.1.G(iii), \( \text{rank}(A) = \text{dim} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) + \text{dim} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right) = \text{rank} \left( \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} \right) \).
**Theorem 3.3.5.** Let $A$ be an $n \times k$ matrix and $B$ be a $k \times m$ matrix. Then $	ext{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$.

**Proof.** Let the columns of $A$ be $a_1, a_2, \ldots, a_k$, the columns of $B$ be $b_1, b_2, \ldots, b_m$, and the columns of $AB$ be $c_1, c_2, \ldots, c_m$. Recall (see the note on page 5 of the class notes for Section 3.2) that if $x \in \mathbb{R}^k$ then $Ax$ is a linear combination of the columns of $A$; that is, $Ax \in \mathcal{V}(A)$. Now from the definition of matrix multiplication, we have $c_i = Ab_i$ for $i = 1, 2, \ldots, m$, so that $c_i = Ab_i \in \mathcal{V}(A)$ for $i = 1, 2, \ldots, m$. So every linear combination of the columns of $AB$ is also a linear combination of the columns of $A$, and $\mathcal{V}(AB)$ is a subspace of $\mathcal{V}(A)$. Hence $	ext{rank}(AB) \leq \text{rank}(A)$. By Theorem 3.2.2, $	ext{rank}(A) = \text{rank}(A^T)$, $	ext{rank}(B) = \text{rank}(B^T)$, and

$$\text{rank}(AB) = \text{rank}((AB)^T).$$

So the previous argument shows that

$$\text{rank}(AB) = \text{rank}((AB)^T) = \text{rank}(B^TA^T) \leq \text{rank}(B^T) = \text{rank}(B).$$

Therefore, $\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$. □

**Theorem 3.3.6.** Let $A$ and $B$ be $n \times m$ matrices. Then

$$|\text{rank}(A) - \text{rank}(B)| \leq \text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B).$$

**Proof.** By Theorem 3.2.2 we have

$$\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_m & 0 \\ I_m & 0 \end{bmatrix} = \begin{bmatrix} AI_m + BI_m & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} A + B & 0 \\ 0 & 0 \end{bmatrix}$$

(or, eliminating the 0 matrices as Gentle does, $[A | B] \begin{bmatrix} I_m \\ I_m \end{bmatrix} = A + B$).

So by Theorem 3.3.5,

$$\text{rank} \left( \begin{bmatrix} A + B & 0 \\ 0 & 0 \end{bmatrix} \right) \leq \min \left\{ \text{rank} \left( \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \right), \text{rank} \left( \begin{bmatrix} I_m & 0 \\ I_m & 0 \end{bmatrix} \right) \right\} \leq \text{rank} \left( \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \right).$$

**Theorem 3.3.6 (continued 1)**

**Proof (continued).** By Theorem 3.3.4(iii),

$$\text{rank} \left( \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \right) \leq \text{rank} \left( \begin{bmatrix} A \\ 0 \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} B \\ 0 \end{bmatrix} \right)$$

and so, combining these last two results,

$$\text{rank} \left( \begin{bmatrix} A + B & 0 \\ 0 & 0 \end{bmatrix} \right) \leq \text{rank} \left( \begin{bmatrix} A \\ 0 \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} B \\ 0 \end{bmatrix} \right).$$

Now the 0 matrices in the second rows of these matrices do not affect ranks. That is, rank $\left( \begin{bmatrix} A + B & 0 \\ 0 & 0 \end{bmatrix} \right) = \text{rank}([A + B | 0])$,

$$\text{rank} \left( \begin{bmatrix} A \\ 0 \end{bmatrix} \right) = \text{rank}(A), \text{ and } \text{rank} \left( \begin{bmatrix} B \\ 0 \end{bmatrix} \right) = \text{rank}(B) \text{ (this can be justified by Theorem 3.3.4(iv) since } \text{rank}(0) = 0).$$

Similarly,

$$\text{rank}([A + B | 0]) = \text{rank}(A + B).$$

Therefore,

$$\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B). \quad (*)$$

**Theorem 3.3.6 (continued 2)**

**Theorem 3.3.6.** Let $A$ and $B$ be $n \times m$ matrices. Then

$$|\text{rank}(A) - \text{rank}(B)| \leq \text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B).$$

**Proof (continued).** With the second inequality established, we have

$$\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B). \quad (*)$$

Next, $A = (A + B) - B$, so by $(*)$ we have

$$\text{rank}(A) = \text{rank}((A + B) - B) \leq \text{rank}(A + B) + \text{rank}(-B)$$

or

$$\text{rank}(A + B) \geq \text{rank}(A) - \text{rank}(-B) = \text{rank}(A) - \text{rank}(B)$$

since $\text{rank}(-B) = \text{rank}(B)$. Similarly (interchanging $A$ and $B$),

$$\text{rank}(A + B) \geq \text{rank}(B) - \text{rank}(A).$$

Therefore,

$$\text{rank}(A + B) \geq |\text{rank}(A) - \text{rank}(B)|.$$
Theorem 3.3.7. Let \( A \) be an \( n \times n \) full rank matrix. Then
\[
(A^{-1})^T = (A^T)^{-1}.
\]

**Proof.** First, \( A^T \) is also \( n \times n \) and full rank by Theorem 3.3.2. We have
\[
A^T (A^{-1})^T = (A^{-1} A)^T \quad \text{by Theorem 3.2.1(1)}
\]
\[
= I^T = I,
\]
so a right inverse of \( A^T \) is \( (A^{-1})^T \). Since \( A \) is full rank and square then, as discussed above, \((A^T)^{-1} = (A^{-1})^T\). □

---

Theorem 3.3.8.

**Theorem 3.3.8.** \( n \times m \) matrix \( A \), where \( n \leq m \), has a right inverse if and only if \( A \) is of full row rank \( n \). \( n \times m \) matrix \( A \), where \( m \leq n \), has a left inverse if and only if \( A \) has full column rank \( m \).

**Proof.** Let \( A \) be an \( n \times m \) matrix where \( n \leq m \) and let \( A \) be of full row rank (that is, \( \text{rank}(A) = n \)). Then the column space of \( A \), \( \mathcal{V}(A) \), is of dimension \( n \) and each \( e_i \), where \( e_i \) is the \( i \)th unit vector in \( \mathbb{R}^n \), is in \( \mathcal{V}(A) \) so that there is \( x_i \in \mathbb{R}^m \) such that \( Ax_i = e_i \) for \( i = 1, 2, \ldots, n \). With \( X \) an \( m \times n \) matrix with columns \( x_i \) and the columns of \( I_n \) as \( e_i \), we have \( AX = I_n \). Also, by Theorem 3.3.6, \( n = \text{rank}(I_n) \leq \min \{ \text{rank}(A), \text{rank}(X) \} \) where \( \text{rank}(A) = n \), so \( \text{rank}(X) = n \) and \( X \) is of full column rank. Furthermore, \( AX = I_n \) has a solution only if \( A \) has full row rank \( n \) since the \( n \) columns of \( I_n \) are linearly independent. That is, \( A \) has a right inverse if and only if \( A \) is of full row rank. The result similarly follows for the left inverse claim. □

---

Theorem 3.3.9.

**Theorem 3.3.9.** If \( A \) is an \( n \times m \) matrix of rank \( r > 0 \) then there are matrices \( P \) and \( Q \), both products of elementary matrices, such that \( PAQ \) is the equivalent canonical form of \( A \), \( PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \).

**Proof.** We prove this by induction. Since \( \text{rank}(A) > 0 \) then some \( a_{ij} \neq 0 \). We move this into position \((1,1)\) by interchanging row 1 and \( i \) and interchanging columns 1 and \( j \) to produce \( E_{1,1} AE_{i,j} \) (we use superscripts of ‘\( c \)’ to denote column operations). Then divide the first row by \( a_{1j} \) to produce an entry of 1 in the \((1,1)\) position (we denote the corresponding elementary matrix as \( E_{1,1} \)) to produce \( B = E_{1,1} AE_{i,j} \). Next we “eliminate” the entries in the first column of \( B \) under the \((1,1)\) entry with the elementary row operations \( R_k \rightarrow R_k - b_{k1} R_1 \) for \( 2 \leq k \leq n \) (we denote the corresponding elementary row matrices as \( E_{k,(-b_{11})1} \) for \( 2 \leq k \leq n \)) to produce
\[
C = E_{n,(-b_{11})1} E_{n-1,(-b_{n-1,1})1} \cdots E_{2,(-b_{21})1} B.
\]

---

Theorem 3.3.9 (continued 1)

**Proof (continued).** Similarly we eliminate the entries in the first row of \( C \) to the right of the \((1,1)\) entry with the elementary column operations \( C_k \rightarrow C_k - c_{1k} C_1 \) (with the corresponding elementary matrices \( E_{n,(-c_{1n})1} \)) to produce
\[
E_{n,(-c_{1n})1} E_{n-1,(-c_{1n})1} \cdots E_{2,(-c_{1n})1} C.
\]

We now have a matrix of the form \( P_1 A Q_1 \) where \( \begin{bmatrix} I_1 & 0_{R_1} \\ 0_{C_1} & X_1 \end{bmatrix} \) where \( R_1 \) is \( 1 \times (n-1) \), \( C_1 \) is \((n-1) \times 1 \), and \( X \) is \((n-1) \times (n-1) \). Also, \( P_1 \) and \( Q_1 \) are products of elementary matrices. By Theorem 3.3.3,
\[
\text{rank}(A) = \text{rank}(P_1 A Q_1) = r. \quad \text{Since} \quad \mathcal{V} \left( \begin{bmatrix} I_1 & 0_{R_1} \\ 0_{C_1} & X_1 \end{bmatrix} \right) \perp \mathcal{V} \left( \begin{bmatrix} 0_{R_1} \\ X_1 \end{bmatrix} \right) \quad \text{then by Theorem 3.3.4(iv)}
\]
\[
r = \text{rank} \left( \begin{bmatrix} I_1 & 0_{R_1} \\ 0_{C_1} & X_1 \end{bmatrix} \right) + \text{rank} \left( \begin{bmatrix} 0_{R_1} \\ X_1 \end{bmatrix} \right) = 1 + \text{rank} \left( \begin{bmatrix} 0_{R_1} \\ X_1 \end{bmatrix} \right) \quad \text{and so}
\]
\[
\text{rank} \left( \begin{bmatrix} 0_{R_1} \\ X_1 \end{bmatrix} \right) = r - 1.
\]
Theorem 3.3.9 (continued 2)

**Proof (continued).** So $\operatorname{rank}(X_1) = r - 1$ (also by Theorem 3.3.4(iv), if you like). If $r - 1 > 0$ then we can similarly find $P_2$ and $Q_2$ products of elementary matrices such that

$$
P_2P_1AQ_1Q_2 = \begin{bmatrix} I_2 & 0_{r, 2} \\ 0_{2, r} & X_2 \end{bmatrix}
$$

and $\operatorname{rank}(X_2) = r - 2$. Continuing this process we can produce

$$
P_rP_{r-1}\cdots P_1AQ_1Q_2\cdots Q_r = \begin{bmatrix} I_r & 0_{r, r} \\ 0_{r, r} & X_r \end{bmatrix}
$$

where $X_r$ has rank 0; that is, where $X_r$ is a matrix of all 0’s. So

$$
P_rP_{r-1}\cdots P_1AQ_1Q_2\cdots Q_r = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}
$$

as claimed.

---

Theorem 3.3.11

**Theorem 3.3.11.** If $A$ is a square full rank matrix (that is, nonsingular) and if $B$ and $C$ are conformable matrices for the multiplications $AB$ and $CA$ then $\operatorname{rank}(AB) = \operatorname{rank}(B)$ and $\operatorname{rank}(CA) = \operatorname{rank}(C)$.

**Proof.** By Theorem 3.3.5,

$$
\operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\} \leq \operatorname{rank}(B).
$$

Also, $B = A^{-1}AB$ so by Theorem 3.3.5, $\operatorname{rank}(B) \leq \min\{\operatorname{rank}(A^{-1}), \operatorname{rank}(AB)\} \leq \operatorname{rank}(AB)$. So $\operatorname{rank}(B) = \operatorname{rank}(AB)$.

Similarly, $\operatorname{rank}(CA) \leq \operatorname{rank}(C)$ and $C = CAA^{-1}$ so $\operatorname{rank}(C) \leq \operatorname{rank}(CA)$ and hence $\operatorname{rank}(C) = \operatorname{rank}(CA)$.

---

Theorem 3.3.12

**Theorem 3.3.12.** If $A$ is a full column rank matrix and $B$ is conformable for the multiplication $AB$, then $\operatorname{rank}(AB) = \operatorname{rank}(B)$. If $A$ is a full row rank matrix and $C$ is conformable for the multiplication $CA$, then $\operatorname{rank}(CA) = \operatorname{rank}(C)$.

**Proof.** Let $A$ be $n \times m$ and of full column rank $m \leq n$. By Theorem 3.3.8, $A$ has a left inverse $A_L^{-1}$ where $A_L^{-1}A = I_m$. By Theorem 3.3.5, $\operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\} \leq \operatorname{rank}(B)$. Now $B = I_mB = A_L^{-1}AB$, so by Theorem 3.3.5 $\operatorname{rank}(B) \leq \min\{\operatorname{rank}(A_L^{-1}), \operatorname{rank}(AB)\} \leq \operatorname{rank}(AB)$, and so $\operatorname{rank}(AB) = \operatorname{rank}(B)$.

Next let $A$ be $n \times m$ and of full column rank $n \leq m$. By Theorem 3.3.8, $A$ has a right inverse $A_R^{-1}$ where $AA_R^{-1} = I_n$. By Theorem 3.3.5, $\operatorname{rank}(CA) = \operatorname{rank}(C)$. Now $C = C I_n = CAA_R^{-1}$, so by Theorem 3.3.5 $\operatorname{rank}(C) \leq \operatorname{rank}(CA)$ and so $\operatorname{rank}(CA) = \operatorname{rank}(C)$.

---

Theorem 3.3.13

**Theorem 3.3.13.** Let $C$ be $n \times n$ and positive definite and let $A$ be $n \times m$.

1. If $C$ is positive definite and $A$ is of full column rank $m \leq n$ then $A^TCA$ is positive definite.

2. If $A^TCA$ is positive definite then $A$ is of full column rank $m \leq n$.

**Proof.** (1) Let $x \in \mathbb{R}^m$, where $x \neq 0$, and let $y = Ax$. So $y$ is a linear combination of the columns of $A$ and since $A$ is of full column rank (so that the columns of $A$ form a basis for the column space of $A$) and $x \neq 0$ then $y \neq 0$. Since $C$ is hypothosized to be positive definite,

$$
x^T(A^TCA)x = (Ax)^TC(Ax) = y^TCy > 0.
$$

Also, $A^TCA$ is $m \times m$ and symmetric since

$$(A^TCA)^T = A^TCT(A^T)^T = A^TCA.$$ Therefore $A^TCA$ is positive definite.
**Theorem 3.3.13.** Let $C$ be an $n \times n$ matrix and let $A$ be an $n \times m$ matrix.

1. If $C$ is positive definite and $A$ is of full column rank $m \leq n$ then $A^T CA$ is positive definite.

2. If $A^T CA$ is positive definite then $A$ is of full column rank $m \leq n$.

**Proof (continued).** Assume not; assume that $A$ is not of full column rank. Then the columns of $A$ are not linearly independent and so with $a_1, a_2, \ldots, a_m$ as the columns of $A$, there are scalars $x_1, x_2, \ldots, x_m$ not all 0, such that $x_1 a_1 + x_2 a_2 + \cdots + x_m a_m = 0$. But then $x \in \mathbb{R}^m$ with entries $x_i$ satisfies $x \neq 0$ and $Ax = 0$. Therefore $x^T (A^T CA)x = (x^T A^T C)(Ax) = (x^T A^T C)0 = 0$, and so $A^T CA$ is not positive definite, a CONTRADICTION. So the assumption that $A$ is not of full column rank is false. Hence, $A$ is of full column rank.

**Theorem 3.3.14.** Properties of $A^T A$.

Let $A$ be an $n \times m$ matrix.

1. $A^T A = 0$ if and only if $A = 0$.

2. $A^T A$ is nonnegative definite.

3. $A^T A$ is positive definite if and only if $A$ is of full column rank.

4. $(A^T A)B = (A^T A)C$ if and only if $AB = AC$, and $B(A^T A) = C(A^T A)$ if and only if $BA^T = CA^T$.

5. $A^T A$ is of full rank if and only if $A$ is of full column rank.

6. rank($A^T A$) = rank($A$).

The product $A^T A$ is called a Gramian matrix.

**Proof.** (1) If $A = 0$ then $A^T = 0$ and $A^T A = 00 = 0$. If $A^T A = 0$ then $tr(A^T A) = 0$. Now the $(i, j)$ entry of $A^T A$ is $\sum_{k=1}^{n} a_{ik}^2 a_{kj} = \sum_{k=1}^{n} a_{ik} a_{kj}$ and so the diagonal $(i, i)$ entry is $\sum_{k=1}^{n} a_{ik}^2$. Hence

$$0 = tr(A^T A) = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ki}^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2$$

(2) For any $y \in \mathbb{R}^m$ we have

$$y^T (A^T A)y = (Ay)^T (Ay) = \|Ay\|^2 \geq 0.$$

(3) From (2), $y^T (A^T A)y = \|Ay\|^2$, so $y^T (A^T A)y = 0$ if and only if $\|Ay\| = 0$. Now $Ay$ is a linear combination of the columns of $A$ so if $A$ is of full column rank then $Ay = 0$ if and only if $y = 0$. That is, if $A$ is of full column rank then for $y \neq 0$ we have $y^T (A^T A)y = \|Ay\|^2 > 0$ and $A^T A$ is positive definite.

If $A$ is not of full column rank then the columns of $A$ are not linearly independent and with $a_1, a_2, \ldots, a_n$ as the columns of $A$, there are scalars $y_1, y_2, \ldots, y_n$ not all 0, such that $y_1 a_1 + y_2 a_2 + \cdots + y_n a_n = 0$. Then the $y \in \mathbb{R}^n$ with entries $y_i$ we have $y \neq 0$ and $Ay = 0$. Then

$$y^T (A^T A)y = \|Ay\|^2 = 0,$$

and so $A^T A$ is not positive definite.

(4) Suppose $A^T AB = A^T AC$. Then

$$A^T AB = A^T AC = 0 \text{ or } A^T A(B-C) = 0,$$

and so

$$(B^T - C^T)A^T A(B-C) = 0.$$

Hence $(A(B-C))^T (A(B-C)) = 0$ and by Part (1), $A(B-C) = 0$. That is, $AB = AC$. Therefore $A^T AB = A^T AC$ if and only if $AB = AC$.

Now suppose $BA^T = CA^T$. Then $BA^T - CA^T = 0$ or $(B-C)A^T A = 0$, and so $(B-C)A^T A(B^T - C^T) = 0$. Hence

$$((B-C)A^T) (B^T - C^T) = 0$$

and by Part (1), $(B-C)^T A^T = 0$. That is, $BA^T = CA^T$. Conversely, if $BA^T = CA^T$ then $BA^T - CA^T = 0$. Therefore $BA^T = CA^T$ if and only if $BA^T = CA^T$.

(5) Suppose $A$ is of full column rank $m \leq n$. Then by Theorem 3.3.12, rank($A^T A$) = rank($A$) = $m$. Since $A^T A$ is $m \times m$, then $A^T A$ is of full rank.
Theorem 3.3.14 (continued 3)

**Proof (continued).** Now suppose $A^T A$ if of full rank $m$. Then by Theorem 3.3.5, $m = \text{rank}(A^T A) \leq \min\{\text{rank}(A^T), \text{rank}(A)\} \leq \text{rank}(A)$, and since $A$ is $n \times m$ then $A$ must be of full column rank $m$.

(6) Let $\text{rank}(A) = r$. If $r = 0$ then $A = 0$ and so $A^T A = 0$ and \( \text{rank}(A) = 0 \) and the claim holds. If $r > 0$, then the columns of $A$ can be permuted so that the first $r$ columns are linearly independent. That is, there is a permutation matrix $Q$ such that $AQ = [A_1 \ A_2]$ where $A_1$ is an $n \times r$ matrix of rank $r$ (and by Theorem 3.3.3, $\text{rank}(AQ) = \text{rank}(A) = r$).

So $A$, is of full column rank and so each column of $A_2$ is in the column space of $A_1$. So there is $r \times (m - r)$ matrix $B$ such that $A_2 = A_1 B$. Then $AQ = [A_1 \ A_2] = [A_1 I_r \ A_1 B] = A_1[I_r \ B]$. Hence

\[(AQ)^T = (A_1[I_r \ B])^T = \begin{bmatrix} I_r & 0 \\ B^T & I_m \end{bmatrix} A_1^T \quad \text{and} \quad (AQ)^T(AQ) = \begin{bmatrix} I_r & 0 \\ B^T & I_m \end{bmatrix} A_1^T A_1[I_r \ B]. \]

Define $T = \begin{bmatrix} I_r & 0 \\ -B^T & I_{m-r} \end{bmatrix}$. Then $T(AQ)^T(AQ)T^T = \begin{bmatrix} I_r & 0 \\ -B^T & I_{m-r} \end{bmatrix} A_1^T A_1[I_r \ B]$. Define $T = \begin{bmatrix} I_r & 0 \\ -B^T & I_{m-r} \end{bmatrix}$.

(7) Now

\[
T(AQ)^T(AQ)T^T = \begin{bmatrix} I_r & 0 \\ -B^T & I_{m-r} \end{bmatrix} A_1^T \begin{bmatrix} I_r & 0 \\ -B^T & I_{m-r} \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ -B^T & I_{m-r} \end{bmatrix} A_1^T = \begin{bmatrix} I_r I_r & 0 \\ -B^T I_r + I_{m-r} B^T \end{bmatrix} A_1^T \]

and

\[
(AQ)^T T = (T(AQ)^T)^T = \begin{bmatrix} A_1^T \ 0 \end{bmatrix}^T = [A_1 \ 0].
\]

Theorem 3.3.14 (continued 4)

**Proof (continued).** Then $T$ is $m \times m$ and of full rank $m$ (as is $T^T$), so by Theorem 3.3.12

\[
\text{rank}(A^T A) = \text{rank}((AQ)^T(AQ)) = \text{rank}(T(AQ)^T(AQ)) = \text{rank}(T(AQ)^T(AQ)T^T). \quad (\ast)
\]

Now

\[
T(AQ)^T = \begin{bmatrix} I_r & 0 \\ -B^T & I_{m-r} \end{bmatrix} \begin{bmatrix} I_r \\ B^T \end{bmatrix} A_1^T = \begin{bmatrix} I_r I_r + 0B^T \\ -B^T I_r + I_{m-r} B^T \end{bmatrix} A_1^T \]

and

\[
(AQ)^T = (T(AQ)^T)^T = \begin{bmatrix} A_1^T \\ 0 \end{bmatrix}^T = [A_1 \ 0].
\]

Theorem 3.3.14 (continued 5)

**Proof (continued).** So

\[
T(AQ)^T(AQ)T^T = \begin{bmatrix} A_1^T & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_1 \ 0 \end{bmatrix} = \begin{bmatrix} A_1^T A_1 & 0 \\ 0 & 0 \end{bmatrix}
\]

(the matrix products are justified by Theorem 3.2.2). So by ($\ast$), \[ \text{rank}(A^T A) = \text{rank} \begin{bmatrix} A_1^T A_1 & 0 \\ 0 & 0 \end{bmatrix} = \text{rank}(A_1^T A_1). \]

Since $A_1$ is of full column rank $r$, by Part (5) $A^T A$ if of full rank $r$. So \[ \text{rank}(A^T A) = \text{rank}(A_1^T A_1) = r = \text{rank}(A), \]

as claimed.

Theorem 3.3.15

**Theorem 3.3.15.** If $A$ is a $n \times n$ matrix and $B$ is $n \times \ell$ then \[ \text{rank}(AB) \geq \text{rank}(A) + \text{rank}(B) - n. \]

**Proof.** Let $r = \text{rank}(A)$. By Theorem 3.3.9, there are $n \times n$ matrices $P$ and $Q$ which are products of elementary matrices such that \[ PAQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}. \]

Let $C = P^{-1} \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix} Q^{-1}$ and then

\[
A + C = P^{-1} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q^{-1} + P^{-1} \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix} Q^{-1} = P^{-1}IQ^{-1} = P^{-1}Q^{-1}.
\]

Now $P^{-1}$ and $Q^{-1}$ are of full rank $n$ (see the notes before the definition of inverse matrix), so by Theorem 3.3.11,

\[
\text{rank}(C) = \text{rank} \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix} = \text{rank}(I_{n-r}) = n - \text{rank}(A). \quad (\ast)
\]
**Theorem 3.3.15.** If $A$ is a $n \times n$ matrix and $B$ is $n \times \ell$ then \( \text{rank}(AB) \geq \text{rank}(A) + \text{rank}(B) - n \).

**Proof (continued).** So for $n \times \ell$ matrix $B$,

\[
\begin{align*}
\text{rank}(B) &= \text{rank}(P^{-1}Q^{-1}B) \text{ by Theorem 3.3.11} \\
&= \text{rank}(AB + CB) \text{ since } A + C = P^{-1}Q^{-1}B \\
&\leq \text{rank}(AB) + \text{rank}(CB) \text{ by Theorem 3.3.6} \\
&\leq \text{rank}(AB) + \text{rank}(C) \text{ by Theorem 3.3.5} \\
&= \text{rank}(AB) + n - \text{rank}(A) \text{ by (*)}.
\end{align*}
\]

So \( \text{rank}(A) + \text{rank}(B) - n \leq \text{rank}(AB) \).

\[\square\]

**Theorem 3.3.16.** $n \times n$ matrix $A$ is invertible if and only if \( \det(A) \neq 0 \).

**Proof.** By Theorem 3.2.4, \( \det(AB) = \det(A)\det(B) \), so if \( A^{-1} \) exists then \( \det(A) = 1/\det(A^{-1}) \) and so \( \det(A) \neq 0 \).

Conversely, if \( \det(A) \neq 0 \) then by Theorem 3.1.3, \( A^{-1} = (1/\det(A))\text{adj}(A) \) and $A$ is invertible.

\[\square\]

**Theorem 3.3.18.** If $A$ and $B$ are $n \times n$ full rank matrices then the Kronecker product satisfies \( (A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \).

**Proof.** Since $A$ and $B$ are full rank, then $A^{-1}$ and $B^{-1}$ exist. Let $A = [a_{ij}]$ and $A^{-1} = [c_{ij}]$. Then \( (A \otimes B)(A^{-1} \otimes B^{-1}) = (a_{ij}B \otimes a_{ij}B)^{-1} \).

\[
\begin{align*}
&= \begin{bmatrix}
a_{11}B & a_{12}B & \cdots & a_{1n}B \\
a_{21}B & a_{22}B & \cdots & a_{2n}B \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1}B & a_{n2}B & \cdots & a_{nn}B
\end{bmatrix}
\begin{bmatrix}
c_{11}B^{-1} & c_{12}B^{-1} & \cdots & c_{1n}B^{-1} \\
c_{21}B^{-1} & c_{22}B^{-1} & \cdots & c_{2n}B^{-1} \\
\vdots & \vdots & \ddots & \vdots \\
c_{n1}B^{-1} & c_{n2}B^{-1} & \cdots & c_{nn}B^{-1}
\end{bmatrix} \\
&= \sum_{k=1}^{n} a_{ik}c_{kj}l_{n} \text{ since } (a_{ik}B)(c_{kj}B^{-1}) = a_{ik}c_{kj}l_{n} \\
&= l_{n^2},
\end{align*}
\]

and so \( A^{-1} \otimes B^{-1} = (A \otimes B)^{-1} \).

\[\square\]