7. Linear Independence and the Rank of a Matrix

The results in this section provide another condition which is necessary and sufficient for the existence of the inverse for a square matrix.

Let \( A = [A_1, A_2, \ldots, A_N] \) be an \( M \times N \) matrix. We say that the \( N \) column vectors of \( A \) are \textit{linearly independent} if the only solution \( x \) to

\[
(22) \quad Ax = 0_M
\]

is \( x = 0_N \). If a solution vector \( x \neq 0_N \) to (22) exists, then we say that the columns of \( A \) are \textit{linearly dependent}.

How can we determine whether the columns of \( A \) are linearly dependent or independent? The Gaussian triangularization algorithm developed in section 3 above can be used to answer this question.

Consider \textit{Stage 1} of the Gaussian algorithm. If we end up in case (iii) (so that \( A_{1*} = 0_M \)), then we can satisfy (22) by choosing \( x = e_1 \) (where \( e_1 \equiv (1, 0_{M-1}^T) \) is the first unit vector of dimension \( M \)). In this case where \( A_{1*} = 0_M \), we can immediately deduce that the columns of \( A \) are linearly dependent.

Now assume that cases (i) or (ii) occurred in Stage 1 of the algorithm and we move on to \textit{Stage 2} (assuming \( N \) and \( M \) are greater than one) of the algorithm. If case (iii) occurs in Stage 2, then at the end of Stage 2, the first two columns of the transformed \( A \) matrix have the following form:

\[
(23) \quad \begin{bmatrix} u_{11} & u_{12} \\ 0_{M-1} & 0_{M-1} \end{bmatrix}
\]

where \( u_{11} \neq 0 \). Consider solving the following equation:

\[
(24) \quad u_{11} x_1 + u_{12} x_2 = 0.
\]

If we set \( x_2 = 1 \), then since \( u_{11} \neq 0 \), we can solve (24) for \( x_1 \) as follows:

\[
(25) \quad x_1 = -u_{12}/u_{11}.
\]

Let the \( M \times M \) matrix \( E \) denote the product of the elementary row operation matrices that transform the first two columns of \( A \) into the case (iii) upper triangular matrix defined by (23). Now premultiply both sides of (22) by \( E \) to obtain:

\[
(26) \quad EA x = E0_M = 0_M.
\]

It can be seen, using (23) - (25), that if we choose \( x \) to be the following vector:
(27) \( x^* = -(u_{12}/u_{11})e_1 + e_2 \neq 0_N, \)

then \( x^* \) satisfies (26). Recall from the previous section that each elementary row matrix that adds a multiple of one row to another row has a determinant equal to one. Since \( E \) is a product of these matrices, its determinant will also equal one. Hence \( E^{-1} \) exists and we can premultiply both sides of \( EAx^* = 0_M \) by \( E^{-1} \) and conclude that \( Ax^* = 0_M \) with \( x^* \neq 0_N \). Thus if case (iii) occurs at the end of Stage 2 of the Gaussian triangularization algorithm, we can conclude that the columns of \( A \) are linearly dependent.

We now need to consider two cases dependent on whether the number of rows of \( A \) (\( M \)) is greater or less than the number of columns of \( A \) (\( N \)).

**Case (1):** \( M \geq N \).

In this case, we follow the Gaussian algorithm through all \( N \) stages. If at the end of any stage (say stage \( i \)) of the algorithm, we find that \( u_{ii} = 0 \), we can adapt the above stage 2 argument to show that there is a nonzero \( x^* \) vector (which has \( x_{i}^* = 1 \) and \( x_{j}^* = 0 \) for \( j > i \)) such that \( Ax^* = 0_M \) and hence the columns of \( A \) are linearly dependent.

On the other hand, if all of the diagonal elements of the final upper triangular matrix are nonzero, then we can show that the columns of \( A \) are linearly independent. In this case, the final \( U \) matrix has the following form:

\[
\begin{pmatrix}
0 & u_{12} & \ldots & u_{1N} \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & u_{NN} \\
0 & 0 & \ldots & 0 \\
\end{pmatrix}
\]

Let \( E \) represent the product of the elementary row matrices that transform \( A \) into the \( U \) defined by (28); i.e., we have

\[
(29) \quad EA = U \quad ; \quad |E| = 1.
\]

Premultiply both sides of (22), \( Ax = 0_M \), by \( E \) to obtain:

\[
(30) \quad EAx = Ux = E0_M = 0_M.
\]

Using (28), we see that the \( N \)th equation in (31) is:

\[
(31) \quad u_{NN} x_N = 0
\]
and since $u_{NN} \neq 0$ by hypothesis, we must have $x_N = 0$. Now look at the N-1st equation in (30):

\[(32) \quad u_{N-1,N-1} x_{N-1} + u_{N-1,N} x = 0.\]

Substituting $x_N = 0$ into (32) yields

\[(33) \quad u_{N-1,N-1} x_{N-1} = 0\]

and since $u_{N-1,N-1} \neq 0$ by hypothesis, we must have $x_{N-1} = 0$. Continuing on in the same way, we deduce that the only x solution to (30) is $x^* = 0_N$.

It is obvious that $x^* = 0_N$ satisfies $Ax = 0_M$. Could there be any other solution to $Ax = 0_M$? Let $x^{**}$ be such that

\[(34) \quad Ax^{**} = 0_M.\]

Premultiplying both sides of (34) by $E$ leads to:

\[(35) \quad EAx^{**} = Ux^{**} = 0_M.\]

But the only solution to (35) is $x^{**} = 0_N$. Hence under our Case (1) hypothesis where $M \geq N$ and all $u_{ii} \neq 0$, $i = 1, 2, \ldots, N$, we deduce that the columns of $A$ are linearly independent. If any of the $u_{ii} = 0$, then the columns of $A$ are linearly dependent.

**Case 2: $M < N$.**

In this case, carry out the Gaussian triangularization procedure until we run out of rows. The final U matrix will have the following form:

\[
U = \begin{bmatrix}
\underline{u_{11}} & u_{12} & \cdots & u_{1M} & u_{1M+1} & \cdots & u_{1N} \\
0 & u_{22} & \cdots & u_{2M} & u_{2M+1} & \cdots & u_{2N} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & u_{MM} & u_{MM+1} & \cdots & u_{MN}
\end{bmatrix}
\]

If any of the $u_{ii} = 0$, then we can adapt our previous arguments to show that the columns of $A$ are linearly dependent. For example, suppose $u_{22}$ is the first zero $u_{ii}$. Then the $x^* \neq 0_N$ defined by (27) will satisfy $Ax^* = 0_M$.

If $u_{ii} \neq 0$ for $i = 1, 2, \ldots, M$, then consider the equations $Ux = 0_M$. If we set $x^*_{M+1} = -1$ and $x^*_{M+2} = x^*_{M+3} = \ldots = x^*_N = 0$, then the equations $Ux = 0_N$ reduce to
which can readily be solved for \( x_1^*, \ldots, x_M^* \); i.e.,

\[
x_M^* = \frac{u_{MM+1}}{u_{MM}};
\]

\[
x_{M-1}^* = \left[\frac{u_{M-1,M+1} x_M^*}{u_{M-1,M-1}}\right];
\]

etc.

The resulting \( x^* \neq 0_N \) and hence we deduce that the columns of \( A \) are linearly dependent.

Thus if we are in Case (2), we inevitably deduce that the columns of \( A \) are linearly dependent.

Putting all of the above material together, we find that the columns of \( A \) are linearly dependent unless \( M \geq N \) and the \( N \) \( u_{ii} \) elements in (28) are all nonzero. Only in this last case, are the columns of \( A \) linearly independent.

**Definition:** The rank of an \( N \) by \( M \) matrix is the maximal number of linearly independent columns which it contains.

Example the rank of

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

is 3, the rank of

\[
\begin{bmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 0 & 2 \\
0 & 0 & 1 & 3
\end{bmatrix}
\]

is also 3.

**Lemma (11):** If the rank of the \( N \) by \( N \) matrix \( A \) is \( N \), then \( A^{-1} \) exists.

**Proof:** If the rank of the \( N \) by \( N \) matrix is \( N \), then all of the columns of \( A \) are linearly independent. Hence, when implementing the Gaussian triangularization of \( A \), all of the diagonal elements \( u_{ii} \) of the upper triangular matrix \( U \) must be nonzero. Hence the determinant of \( U = \prod_{i=1}^{N} u_{ii} \) is also nonzero. Recall that

\[
EA = U \quad \text{where} \quad |E| = 1.
\]

Hence, taking determinants on both sides of (39):

\[
|EA| = |E| |A| = |A| |U| = \prod_{i=1}^{N} u_{ii} \neq 0,
\]
and we conclude that \( |A| \neq 0 \) so \( A^{-1} \) exists. Q.E.D.

**Problem 14:** Let \( A \) be \( M \times N \) where \( M > N \) and consider the system of equations

\[
(i) \quad Ax = b
\]

where \( x \) is an \( N \) dimensional solution vector and \( b \) is an \( M \) dimensional vector of parameters. Suppose the \( N \) columns of \( A = [A_{\bullet 1}, A_{\bullet 2}, \ldots, A_{\bullet M}] \) are linearly independent. Under what conditions on \( b \) will a solution \( x \) to (i) exist and how could you compute it if it did exist? *Hint:* Make use of the \( M \times M \) elementary row matrix \( E \) which reduces \( A \) to upper triangular form \( U \); i.e., \( E \) and \( U \) satisfy (28) and (29) in the text above.

8. **Comparative Statics Analysis of a System of Linear Equations**

Let \( A \) be an \( N \times N \) matrix and \( b \) an \( N \) dimensional vector. If \( |A| \neq 0 \), then the solution \( x \) to \( Ax = b \) can be written as:

\[
(41) \quad x = A^{-1}b.
\]

Obviously, the components of the solution vector \( x \) depend on the components \( a_{ij} \) of \( A \) and \( b_i \) of \( b = [b_1, b_2, \ldots, b_N]^T \). How does \( x \) change as the \( a_{ij} \) and \( b_i \) change?

Using (41), the \( N \times N \) matrix of the derivatives of \( x_i \) with respect to \( b_j \), \( \partial x_i / \partial b_j \), can be written as

\[
(42) \quad d(bx) = [\partial x_i / \partial b_j] = A^{-1}.
\]

Recalling the determinantal formula for \( A^{-1} \) given in Lemma (11), we see that

\[
(43) \quad \partial x_i / \partial b_j = A_{ji} / |A| ; \quad 1 \leq i, j \leq N
\]

where \( A_{ji} \) is the \( j \)th cofactor of \( A \).

In order to determine how the components of \( x \) change as the components of \( A \) change, it is convenient to study a somewhat more general problem: we let each component of the \( A \) matrix be a function of the scalar variable \( t \) (i.e., \( a_{ij} = a_{ij}(t) \) for \( 1 \leq i, j \leq N \)) and then \( x \) defined by (41) will also be a function of \( t \), \( x(t) \). We then compute the vector of derivatives, \( x'(t) = [x_1'(t), \ldots, x_N'(t)]^T \). Before we do this, we establish a preliminary result.

**Lemma (12):** Let \( A(t) = [a_{ij}(t)] \) have \( N \) columns and \( B(t) = [b_{ij}(t)] \) have \( N \) rows so that \( C(t) = A(t)B(t) \) is well defined. Note that each element of \( A(t) \) and each
element of \( B(t) \) is a function of the scalar variable \( t \). Then the matrix of derivatives with respect to \( t \) of the product matrix is

\[
\frac{d}{dt} \mathbf{C}(t) = \mathbf{A}(t) \mathbf{B}(t) + \mathbf{A}(t) \frac{d}{dt} \mathbf{B}(t)
\]

where \( \mathbf{A}(t) = \begin{bmatrix} a_{ij}(t) \end{bmatrix} \) and \( \mathbf{B}(t) = \begin{bmatrix} b_{ij}(t) \end{bmatrix} \) are the matrices of derivatives of \( \mathbf{A}(t) \) and \( \mathbf{B}(t) \).

**Proof:** The \( ij \)th element of \( \mathbf{C} \) is

\[
\begin{align*}
c_{ij}(t) = & \left( \mathbf{A} \right)_{i} \left( \mathbf{B} \right)_{j}(t) = \sum_{n=1}^{N} a_{in}(t)b_{nj}(t). \\
\end{align*}
\]

Differentiating (45) with respect to \( t \) yields for all \( i \) and \( j \):

\[
\frac{d}{dt} c_{ij}(t) = \sum_{n=1}^{N} a_{in}(t)\frac{d}{dt}b_{nj}(t) + \sum_{n=1}^{N} a_{in}(t)b_{nj}(t).
\]

It can be seen that equations (46) are equivalent to equations (44).

Q.E.D.

Now let the \( \mathbf{B}(t) \) matrix which appears in Lemma (12) be \( \mathbf{A}^{-1}(t) \) and differentiate both sides of the following identity with respect to \( t \):

\[
\frac{d}{dt} \mathbf{A}(t) \mathbf{A}^{-1}(t) = \mathbf{I}_N.
\]

Using Lemma (12), we obtain:

\[
\mathbf{A}(t)\frac{d}{dt} \mathbf{A}^{-1}(t) + \mathbf{A}(t)[\frac{d\mathbf{A}}{dt}]\mathbf{A}^{-1}(t) = 0_{N\times N}
\]

where \( \frac{d}{dt} \mathbf{A}^{-1}(t) = \begin{bmatrix} \frac{d}{dt} a_{ij}^{-1}(t) \end{bmatrix} \) is the \( N \) by \( N \) matrix of derivatives of the components of \( \mathbf{A}^{-1} \) with respect to \( t \). Premultiply both sides of (48) by \( \mathbf{A}^{-1}(t) \) and after rearranging terms, we obtain the following formula:

\[
\frac{d\mathbf{A}^{-1}(t)}{dt} = -\mathbf{A}^{-1}(t)\frac{d\mathbf{A}}{dt}\mathbf{A}^{-1}(t).
\]

Now return to (41) which we rewrite as:

\[
x(t) = \mathbf{A}^{-1}(t)b.
\]

Differentiating (50) with respect to \( t \) and using (49) yields:

\[
\frac{d}{dt} x(t) = \left[ \frac{d}{dt} x(t) \right] = \left[ \begin{array}{c} \frac{d}{dt} x_1(t) \\ \vdots \\ \frac{d}{dt} x_N(t) \end{array} \right] = \begin{bmatrix} a_{ij}(t) \end{bmatrix} \begin{bmatrix} \frac{d}{dt} x_1(t) \\ \vdots \\ \frac{d}{dt} x_N(t) \end{bmatrix} = \begin{bmatrix} a_{ij} \end{bmatrix} \begin{bmatrix} \frac{d}{dt} x_1(t) \\ \vdots \\ \frac{d}{dt} x_N(t) \end{bmatrix}.
\]

If only \( a_{ij} \) depends on \( t \), then
\[(52) \quad A(t) = e_i e_j^T a_{ij}(t)\]

where \(e_i\) and \(e_j\) are the \(i\) and \(j\)th unit vectors. Substituting (52) into (51) yields in this special case:

\[(53) \quad x(t) = [A^{\text{T}}(t)e_i e_j^T a_{ij}(t)] A^{\text{T}}(t)b.\]

**Problem 15:** Suppose the \(N\) components of the \(b\) vector are all functions of the scalar variable \(t\); i.e., we have \(b(t) = [b_1(t), \ldots, b_N(t)]^T\). Define

\[(i) \quad x(t) = A^{-1} b(t)\]

where the \(N\) by \(N\) matrix \(A\) does not depend on \(t\) and \(|A| \neq 0\). Exhibit a formula for the vector of derivatives \(x(t)\). *Hint:* This problem is easy!

**Problem 16:** Let \(A\) be an \(N\) by \(N\) matrix. Regard \(|A|\) as a function of the \(ij\)th element of \(A\), \(a_{ij}\); i.e., define the function \(f(a_{ij}) = |A|\). Find a formula for the derivatives of the determinant of \(A\) with respect to \(a_{ij}\); i.e., calculate \(f'(a_{ij})\). *Hint:* use Lemma (10).