Linear Dependence and Independence

While short in length, the following material on linear dependence and independence is of fundamental importance—so much so that it forms a separate chapter.

3.1 Definitions

Any finite set of row or column vectors, or more generally any finite set of matrices, is either linearly dependent or linearly independent. A nonempty finite set \( \{A_1, A_2, \ldots, A_k\} \) of \( m \times n \) matrices is said to be linearly dependent if there exist scalars \( x_1, x_2, \ldots, x_k \), not all zero, such that
\[
x_1 A_1 + x_2 A_2 + \cdots + x_k A_k = 0.
\]
If no such scalars exist, the set is called linearly independent. The empty set is considered to be linearly independent. Note that if any subset of a finite set of matrices is linearly dependent, then the set itself is linearly dependent.

While technically linear dependence and independence are properties of sets of matrices, it is customary to speak of “a set of linearly dependent (or independent) matrices” or simply of “linearly dependent (or independent) matrices” instead of “a linearly dependent (or independent) set of matrices.” In particular, in the case of row or column vectors, it is customary to speak of “linearly dependent (or independent) vectors.”

3.2 Some Basic Results

Note that a set consisting of a single matrix is linearly dependent if the matrix is null, and is linearly independent if the matrix is nonnull. An elementary result on the linear dependence or independence of two or more matrices is expressed in the following lemma.
Lemma 3.2.1. A set \( \{A_1, A_2, \ldots, A_k\} \) of two or more \( m \times n \) matrices is linearly dependent if and only if at least one of the matrices is expressible as a linear combination of the others, that is, if and only if, for some integer \( j \) \( (1 \leq j \leq k) \), \( A_j \) is expressible as a linear combination of \( A_1, \ldots, A_{j-1}, A_{j+1}, \ldots, A_k \).

**Proof.** Suppose that, for some \( j \), \( A_j \) is expressible as a linear combination
\[
A_j = x_1A_1 + \cdots + x_{j-1}A_{j-1} + x_{j+1}A_{j+1} + \cdots + x_kA_k
\]
of the other \( k-1 \) \( m \times n \) matrices. Then,
\[
(-x_1)A_1 + \cdots + (-x_{j-1})A_{j-1} + A_j + (-x_{j+1})A_{j+1} + \cdots + (-x_k)A_k = 0,
\]
implying that \( \{A_1, A_2, \ldots, A_k\} \) is a linearly dependent set.

Conversely, suppose that \( \{A_1, A_2, \ldots, A_k\} \) is linearly dependent, in which case
\[
x_1A_1 + x_2A_2 + \cdots + x_kA_k = 0
\]
for some scalars \( x_1, x_2, \ldots, x_k \), not all zero. Let \( j \) represent any integer for which \( x_j \neq 0 \). Then,
\[
A_j = (-x_1/x_j)A_1 + \cdots + (-x_{j-1}/x_j)A_{j-1} + (-x_{j+1}/x_j)A_{j+1} + \cdots + (-x_k/x_j)A_k.
\]
Q.E.D.

Another, more profound result on the linear dependence or independence of two or more matrices is expressed in the following lemma.

Lemma 3.2.2. A set \( \{A_1, A_2, \ldots, A_k\} \) of two or more \( m \times n \) matrices, the first of which is nonnull, is linearly dependent if and only if at least one of the matrices is expressible as a linear combination of the preceding ones, that is, if and only if, for some integer \( j \) \( (2 \leq j \leq k) \), \( A_j \) is expressible as a linear combination of \( A_1, \ldots, A_{j-1} \). (Or, equivalently, \( \{A_1, A_2, \ldots, A_k\} \) is linearly independent if and only if none of the matrices is expressible as a linear combination of its predecessors, that is, if and only if, for every integer \( j \) \( (2 \leq j \leq k) \), \( A_j \) is not expressible as a linear combination of \( A_1, \ldots, A_{j-1} \).)

**Proof.** Suppose that, for some \( j \), \( A_j \) is expressible as a linear combination of \( A_1, \ldots, A_{j-1} \). Then, it follows immediately from Lemma 3.2.1 that the set \( \{A_1, \ldots, A_k\} \) is linearly dependent.

Conversely, suppose that \( \{A_1, \ldots, A_k\} \) is linearly dependent. Define \( j \) to be the first integer for which \( \{A_1, \ldots, A_j\} \) is a linearly dependent set. Then,
\[
x_1A_1 + \cdots + x_jA_j = 0
\]
for some scalars \( x_1, \ldots, x_j \), not all zero. Moreover, \( x_j \neq 0 \), since otherwise \( \{A_1, \ldots, A_{j-1}\} \) would be a linearly dependent set, contrary to the definition of \( j \). Thus,