The Proof of Lemma 4.11 (pp.270–271)

Lemma 4.11 states:

**Lemma 1** Let $V$ be a vector space, and let $S$ and $T$ be subsets of $V$. If

(a) $S$ is a finite set,

(b) $\text{span}(S) = V$, and

(c) $T$ is a linearly independent set,

then $T$ is a finite set, and $|T| \leq |S|$.

The proof of this lemma may well be the hardest you have ever encountered. It is certainly the hardest you have seen so far in this course: the calculations needed for the proof are not easy to follow, and it far from obvious how anyone thought to write them down in the first place. This handout is my attempt to guide you through the proof. First, I will outline the entire argument; then, I will introduce some useful notation; and finally, I will focus in on the difficult part of the proof.

**The Outline of the Proof.**

We first dispose of the case $S = \emptyset$: in this case,

$$\text{span}(S) = \text{span} \emptyset = \{0_V\},$$

and since the only linearly independent subset of the vector space $V = \{0_V\}$ is the set $T = \emptyset$, we have that necessarily

$$|T| = 0 = |S|.$$

We then move to the case $S \neq \emptyset$; let $S = (v_1, \ldots, v_n)$, where $n \geq 1$. Suppose, to get a contradiction, that the conclusion is false, so that $T$ is either an infinite set or a finite set such that $|S| < |T|$. In either case, there would be a linearly independent sequence

$$(w_1, \ldots, w_n, w_{n+1})$$

Generally, a lemma is a difficult theorem whose importance is the consequences one can derive from it. I believe I know how, but explaining my answer would involve introducing a different (and at least equally difficult!) proof of the lemma.
of vectors chosen from $T$. The rest of the proof—this is the hard part—establishes that then necessarily

$$w_{n+1} \in \text{span}(w_1, \ldots, w_n),$$

which means that $(w_1, \ldots, w_n, w_{n+1})$ would be linearly dependent (as well as being linearly independent), contradiction.
A Preliminary Step: Some Useful Notation.

It will be useful to write linear combinations so as to make them look like matrix products: for a sequence of vectors \((r_1, \ldots, r_t)\) chosen from \(V\) and a sequence of scalars \((a_1, \ldots, a_t)\),

\[
\text{the linear combination } a_1 r_1 + \cdots + a_t r_t \text{ will be written } \begin{bmatrix} a_1 & \cdots & a_t \end{bmatrix} \cdot \begin{bmatrix} r_1 \\ \vdots \\ r_t \end{bmatrix}.
\]

(1)

In the case that \(V = \mathbb{R}^m\), one can prove—not just define—the expressions in (1) to be equal:

**Exercise 1** Suppose that in equation (1), the vectors in the sequence \((r_1, \ldots, r_t)\) are chosen from the vector space \(V = \mathbb{R}^m\). Prove: if you write the vectors \(r_1, \cdots, r_t\) with coordinates, so that the expression on the right side of (1) becomes an ordinary matrix product, then the matrix product in (1) and the linear combination in (1) in fact turn out to be the same vector in \(\mathbb{R}^m\).

I will need to use (1) in two places.

1. I will rewrite the equation at the top of p. 258:

\[
w_{n+1} = a_1 v_1 + \cdots + a_n v_n \implies w_{n+1} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.
\]

(2)

2. I will also rewrite the system of equations directly below (2):

\[
w_1 = c_{11} v_1 + \cdots + c_{1n} v_n \\
\vdots \quad \vdots \quad \vdots \quad \vdots \\
w_n = c_{n1} v_1 + \cdots + c_{nn} v_n \\
\implies \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \vdots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}
\]

(3)

Note that the text puts \(C := \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \vdots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}\). This makes it possible to express (3) more simply:

\[
w_1 = c_{11} v_1 + \cdots + c_{1n} v_n \\
\vdots \quad \vdots \quad \vdots \quad \vdots \\
w_n = c_{n1} v_1 + \cdots + c_{nn} v_n \\
\implies \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = C \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.
\]

(4)

---

\(^3\)In each case, the “product” will feature one or more rows of scalars for the left factor and, for the right factor, a column “vector” whose entries are themselves vectors.

\(^4\)This restriction turns out to be unnecessary; you will see why when you study §4.7.
The hard part of the proof turns on knowing that (1) represents an associative operation—in particular, we will need equations (8) and (11) below. The calculations on p.258 and p.259 are in essence proofs that equations (8) and (11) are correct.

The Hard Part of the Proof.

Since \( \text{span}(v_1, \ldots, v_n) = \mathcal{V} \), we know that there is a matrix \( [a_1 \quad \cdots \quad a_n] \) that satisfies (2) as well as a matrix
\[
\begin{bmatrix}
c_{11} & \cdots & c_{1n} \\
\vdots & \ddots & \vdots \\
c_{n1} & \cdots & c_{nn}
\end{bmatrix}
\]
that satisfies (3)/(4).
The first order of business is to establish the following claim.

Claim: the matrix $C$ is invertible.

I will prove the claim by showing that the homogeneous system

$$
C^T \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5}
$$

or equivalently

$$
[x_1 \cdots x_n] \cdot C = [0 \cdots 0] \tag{6}
$$

has only the trivial solution.

Suppose that $u = [u_1 \cdots u_n]$ satisfies (6).\(^6\) Then

$$
u_1 w_1 + \cdots + u_n w_n = [u_1 \cdots u_n] \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = [u_1 \cdots u_n] \cdot \left( C \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \right). \tag{7}
$$

Now, by the calculation on p.258, we have the associativity identity

$$
[u_1 \cdots u_n] \cdot \left( C \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \right) = \left( [u_1 \cdots u_n] \cdot C \right) \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}, \tag{8}
$$

and (8) allows us to complete the calculation begun in (7):

$$
u_1 w_1 + \cdots + u_n w_n =
$$

from (7) $\longrightarrow$ $\begin{bmatrix} u_1 \cdots u_n \end{bmatrix} \cdot \left( C \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \right)$

by (8) $\longrightarrow$ $\left( [u_1 \cdots u_n] \cdot C \right) \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$

by (6) $\longrightarrow$ $[0 \cdots 0] \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$

by (1) $\longrightarrow$ $0v_1 + \cdots + 0v_n$

$\longrightarrow 0v.$

---

\(^5\)The transpose $C^T$ arises because (5) is the standard way of writing a homogeneous system, but the proof will require the system to be written in form (6).

\(^6\)To prove the claim, WTS: $u_1 = u_2 = \cdots = u_n = 0.$
Since \((w_1, \ldots, w_n)\) is linearly independent, it follows that \(u_1 = u_2 = \cdots = u_n = 0. \) (claim)

We now consider the sequence of scalars \((a_1, \ldots, a_n)\) from (2). Since \(C\) and \(C^T\) are invertible, there is a (unique) sequence of scalars \((b_1, \ldots, b_n)\) such that

\[
C^T \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}
\]

or equivalently

\[
\begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix} \cdot C = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}. \quad (9)
\]
We can then substitute (9) into (2):

\[ w_{n+1} = a_1 v_1 + \cdots + a_n v_n = [a_1 \cdots a_n] \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \left( [b_1 \cdots b_n] \cdot C \right) \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}. \] (10)

Now, by the calculation on p.259, we have the associativity identity

\[ \left( [b_1 \cdots b_n] \cdot C \right) \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = [b_1 \cdots b_n] \cdot \left( C \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \right), \] (11)

and (11) allows us to complete the calculation begun in (10):

\[ w_{n+1} = a_1 v_1 + \cdots + a_n v_n = [a_1 \cdots a_n] \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \]

from (10) \(\longrightarrow\) \[ \left( [b_1 \cdots b_n] \cdot C \right) \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \]

by (11) \(\longrightarrow\) \[ [b_1 \cdots b_n] \cdot \left( C \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \right) \]

by (4) \(\longrightarrow\) \[ [b_1 \cdots b_n] \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \]

by (1) \(\longrightarrow\) \[ b_1 w_1 + \cdots + b_n w_n. \]

Thus,

\[ w_{n+1} \in \text{span}(w_1, \ldots, w_n). \]

The hard part of the proof is now complete. \(\blacksquare\)