Math 34600 L (43597) - Lectures 02

Ethan Akin
Office: Marshak 325
Email: eakin@ccny.cuny.edu

Spring, 2024

Contents

Vector Spaces - Introduction

You should recall the definition of a vector as an object with magnitude and direction (as opposed to a scalar with magnitude alone). This was devised as a means of representing forces and velocities which exhibit these vector characteristics. The behavior of these physical phenomena leads to a geometric notion of addition of vectors (the "parallelogram law") as well as multiplication by scalars - real numbers. By using coordinates one discovers important properties of these operations.

Addition Properties

$$
\begin{aligned}
(\mathbf{v} + \mathbf{w}) + \mathbf{z} &= \mathbf{v} + (\mathbf{w} + \mathbf{z}). \\
\mathbf{v} + \mathbf{w} &= \mathbf{w} + \mathbf{v}. \\
\mathbf{v} + \mathbf{0} &= \mathbf{v}. \\
\mathbf{v} + -\mathbf{v} &= \mathbf{0}.
\end{aligned}
\tag{4.1}
$$

The first two of these, the Associative and Commutative Laws, allow us to rearrange the order of addition and to be sloppy about parentheses just as we are with numbers. We use them without thinking about them. The third says that the zero vector behaves the way the zero number does: adding it leaves the vector unchanged. The last says that every vector has an "additive inverse" which *cancels* it. The subtraction $\mathbf{w} - \mathbf{v}$ means $\mathbf{w} + (-\mathbf{v})$.

Scalar Multiplication Properties

$$\begin{aligned} a(b\mathbf{v}) &= (ab)\mathbf{v}. \\ 1\mathbf{v} &= \mathbf{v}. \end{aligned} \tag{4.2}$$

The first of these looks like the Associative Law but it isn't. It relates multiplication between numbers and scalar multiplication. The second says that multiplication by 1 leaves the vector unchanged the way the addition with $\mathbf{0}$ does.

Distributive Properties

$$\begin{aligned} a(\mathbf{v} + \mathbf{w}) &= a\mathbf{v} + a\mathbf{w}. \\ (a + b)\mathbf{v} &= a\mathbf{v} + b\mathbf{v}. \end{aligned} \tag{4.3}$$

These link addition and scalar multiplication.

In Linear Algebra we abstract from the original physical examples. Now a vector is just something in a vector space. A vector space is a set V of objects which are called vectors. What is special about a vector space is that it carries a definition of addition between vectors and a definition of multiplication of a vector by a scalar. For us a scalar is a real number, or possibly a complex number. We discard the original geometric picture and keep just the properties given above. These become the *axioms of a vector space*.

In particular, the zero vector $\mathbf{0}$ is the vector which "zeroes". That is, it satisfies the property given by the axiom $\mathbf{x} + \mathbf{0} = \mathbf{x}$ for all $\mathbf{x} \in V$. There is only one zero vector. In fact, suppose that $\mathbf{v} + \mathbf{q} = \mathbf{v}$ for some vector $\mathbf{v}$. Then $\mathbf{q}$ is a candidate for the zero vector. We show that $\mathbf{q} = \mathbf{0}$ by using the axioms.

1- $\mathbf{q} + \mathbf{v} = \mathbf{v} + \mathbf{q} = \mathbf{v}$. (Commutative Law)

2- $\mathbf{q} + (\mathbf{v} + -\mathbf{v}) = (\mathbf{q} + \mathbf{v}) + -\mathbf{v} = \mathbf{v} + -\mathbf{v}$. (Associative Law)

3- $\mathbf{v} + -\mathbf{v} = \mathbf{0}$. (Cancellation)

5- So $\mathbf{q} + \mathbf{0} = \mathbf{0}$.

5- $\mathbf{q} = \mathbf{q} + \mathbf{0}$. (Zero Property)

6- Therefore $\mathbf{q} = \mathbf{q} + \mathbf{0} = \mathbf{0}$.

This illustrates the use of the axioms in detail. In practise, we use the axioms unconsciously just as we do with numbers. From $\mathbf{v} + \mathbf{q} = \mathbf{v}$, we add $-\mathbf{v}$ to both sides to cancel $\mathbf{v}$, getting $\mathbf{0} + \mathbf{q} = \mathbf{0}$. Since $\mathbf{0} + \mathbf{q} = \mathbf{q}$, it follows that $\mathbf{q} = \mathbf{0}$. We will call this result *The Uniqueness of* $\mathbf{0}$.

There is similarly *The Uniqueness of* $-\mathbf{x}$. Suppose $\mathbf{p}$ cancels $\mathbf{v}$. That is, $\mathbf{v} + \mathbf{p} = \mathbf{0}$. Add $-\mathbf{v}$ to both sides to get $\mathbf{0} + \mathbf{p} = -\mathbf{v}$. Since $\mathbf{p} = \mathbf{0} + \mathbf{p}$, we have $\mathbf{p} = -\mathbf{v}$.

From the it is possible to derive various other simple, but important properties. (Theorem 6.1.3)

$$\begin{array}{rcl} a\mathbf{v} = \mathbf{0} & \iff & a = 0 \quad \text{or} \quad \mathbf{v} = \mathbf{0}. \\ & (-1)\mathbf{v} = -\mathbf{v}. & \end{array} \tag{4.4}$$

Proof: $1\mathbf{v} + 0\mathbf{v} = (1+0)\mathbf{v} = 1\mathbf{v}$. By Uniqueness of $\mathbf{0}$, $0\mathbf{v} = \mathbf{0}$.

$a\mathbf{0} + a\mathbf{0} = a(\mathbf{0} + \mathbf{0}) = a\mathbf{0}$. By Uniqueness of $\mathbf{0}$, $a\mathbf{0} = \mathbf{0}$.

If $a\mathbf{v} = \mathbf{0}$ and $a \neq 0$, then $\mathbf{v} = 1\mathbf{v} = (1/a)a\mathbf{v} = (1/a)\mathbf{0} = \mathbf{0}$.

$\mathbf{0} = 0\mathbf{v} = (1 + (-1))\mathbf{v} = 1\mathbf{v} + (-1)\mathbf{v} = \mathbf{v} + (-1)\mathbf{v}$. By Uniqueness of $-\mathbf{x}$, $(-1)\mathbf{v} = -\mathbf{v}$.

$$\begin{array}{lrcl} & \mathbf{v} + \mathbf{x} = \mathbf{w} & \iff & \mathbf{x} = \mathbf{w} - \mathbf{v}. \\ \text{So} & \mathbf{v} + \mathbf{x}_1 = \mathbf{v} + \mathbf{x}_2 & \implies & \mathbf{x}_1 = \mathbf{x}_2. \end{array} \tag{4.5}$$

Fundamental Principle of Linear Algebra

$$A = B \iff A - B = 0. \tag{4.6}$$

The most important examples of vector spaces are:

- $\mathbf{M}_{mn}$ the set of $m \times n$ matrices with matrix addition and scalar multiplication. In particular, $\mathbb{R}^n$ thought of as column vectors or row vectors.
- The set of all real-valued functions on a set S with $(f+g)(s) =_{def} f(s) + g(s)$ and $(af)(s) =_{def} a(f(s))$.

Let us look at Exercises 6.1/3, 8, page 335.

Subspaces and Linear Combinations

Because any vector space contains its zero vector, it is never an empty set.

A *subspace* W of a vector space V is a nonempty subset which is *closed under vector addition and scalar multiplication*. That is, $\mathbf{v}, \mathbf{w} \in W$ imply $\mathbf{v} + \mathbf{w} \in W$ and $a\mathbf{v} \in W$ for every scalar a. In particular, $\mathbf{0} = 0\mathbf{v} \in W$ and if $\mathbf{v} \in W$ then $-\mathbf{v} = (-1)\mathbf{v} \in W$.

So a subspace W is itself a vector space with the other axioms for W following from the fact that they are true for V.

To check whether W is a subspace, always look first to see if $\mathbf{0} \in W$.

The intersection of any collection of subspaces is a subspace, but the union of two subspaces is usually not a subspace.

Linear Combinations

For a list $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in V and a list $\{c_1, c_2, \ldots, c_k\}$ of scalars,

$$\sum_{i=1}^{k} c_i\mathbf{v}_i = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$$

is the linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ with coefficients $\{c_1, c_2, \ldots, c_k\}$.

From the axioms we can "clear parentheses" to get:

$$\begin{aligned} &a(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k) = \\ &ac_1\mathbf{v}_1 + ac_2\mathbf{v}_2 + \cdots + ac_k\mathbf{v}_k. \end{aligned} \tag{4.7}$$

The scalar multiple of a linear combination for a list is a linear combination for the same list.

We can add linear combinations to get:

$$\begin{array}{cccccccc} & c_1\mathbf{v}_1 & + & c_2\mathbf{v}_2 & + & \dots & + & c_k\mathbf{v}_k \\ + & d_1\mathbf{v}_1 & + & d_2\mathbf{v}_2 & + & \dots & + & d_k\mathbf{v}_k \\ \hline \end{array} \tag{4.8}$$
$$(c_1+d_1)\mathbf{v}_1 + (c_2+d_2)\mathbf{v}_2 + \cdots + (c_k+d_k)\mathbf{v}_k.$$

The sum of two linear combinations for a list is a linear combination for the same list.

Notice that for lists $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_\ell\}$.

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \;=\; c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k + 0\mathbf{w}_1 + \dots 0\mathbf{w}_\ell$$

A linear combination for a list is a linear combination for any larger list. Thus, the sum of linear combinations on two different lists is a linear combination on the combined list.

Let S be a set of vectors in V. A *linear combination on* S is a linear combination for a list $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ of vectors in V. is a vector $\mathbf{v} = \Sigma_{i=1}^{n} c_i \mathbf{v}_i$. By convention, we regard $\mathbf{0}$ to be the linear combination on the empty list of vectors.

Define $span(S)$ to be the set of all linear combinations on S. That is, $\mathbf{v} \in span(S)$ if and only if $\mathbf{v} = \Sigma_{i=1}^{n} c_i \mathbf{v}_i$ for some list $\{\mathbf{v}_1, ..., \mathbf{v}_k\}$ of vectors in V and list $\{c_1, ..., c_k\}$ of scalars.

By the above convention $span(\emptyset) = \{\mathbf{0}\}$.

Theorem 4.01: For S a subset of a vector space V.

(a) If $\mathbf{v} \in S$, then $\mathbf{v}$ is a linear combination on S. That is, $S \subset span(S)$.

(b) If S is contained in a subspace W then every linear combination on S is contained in W. That is, $span(S) \subset W$.

(c) $span(S)$ is a subspace.

(d) $span(S)$ is the intersection of all the subspaces which contain S. That is,

$$span(S) \ = \bigcap \{ \ W : S \subset W \text{ and } W \ \text{ is a subspace of } V \ \}.$$

Proof: (a) $\mathbf{v} = 1\mathbf{v}$ and so $\mathbf{v}$ is a linear combination for the list $\{\mathbf{v}\}$.

(b) Obvious since W is closed under addition and scalar multiplication.

(c) We saw above that a scalar multiple of a linear combination for a list is a linear combination for the same list and the sum of linear combinations for two lists is a linear combination for the combined list.

(d) From (b) $span(S)$ is contained in the intersection. Since by (c) it is itself a subspace it is one of the W's and so contains the intersection.
$\square$

Corollary 4.02: (i) If W is a subspace and $S \subset W$, then $span(S) \subset W$.

(ii) If $S \subset T$, then $span(S) \subset span(T)$.

(iii) If $T \subset span(S)$, then $span(S) = span(S \cup T)$.

Proof: (i) This is (b) of the Theorem.

(ii)By (a) $S \subset T \subset span(T)$ and by (c) $span(T)$ is a subspace. So $span(S) \subset span(T)$ by (i).

(iii) $S \subset S \cup T$ and so $span(S) \subset span(S \cup T)$ by (ii).
By assumption and (a), $S \cup T \subset span(S)$ and by (c) $span(S)$ is a subspace.
So, $span(S \cup T) \subset span(S)$ by (i).
$\square$

We say that S *spans* V, or just S spans, when *span*(S) is the whole space, V. That is, every vector in V is a linear combination on S.

BE ABLE TO DEFINE: (1) W is a subspace of V.

(2) $\mathbf{v}$ is linear combination on $S = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$.

(3) The span of S.

(4) S spans the vector space V.

Let us look at Exercises 6.2/ 2df, 3dg, pages 241, 242.

We can regard an $m \times n$ matrix A as a list of n columns each of length m: $A = (C_1(A)\ C_2(A)\ \dots\ C_n(A))$ or just $(C_1\ C_2\ \dots\ C_n)$ For X an $n \times 1$ matrix $AX \ = \ x_1 C_1 + x_2 C_2 + \dots x_n C_n$. That is, AX is the linear combination of the columns of A with coefficients given by X.

The *Column Space* of A, denoted $COL(A)$ is the span of the columns of A so that

$$COL(A) \ = \ span(\{C_1, \dots, C_n\}) \ = \ \{AX : X \in \mathbb{R}^n\} \quad (4.9)$$

is a subspace of $\mathbb{R}^m$.

Thus, $B \in COL(A)$ meaning B can be written as a linear combination of the columns of A if and only if the system $AX = B$ has a solution, i.e. if and only if it is consistent.

Theorem 4.03: Let A be an $m \times n$ matrix with columns $C_1, \ldots C_n \in \mathbb{R}^m$. Assume that Q in Reduced Echelon Form is row equivalent to A. The following conditions are equivalent.

(i) $COL(A) = \mathbb{R}^m$.

(ii) The set of columns of A spans $\mathbb{R}^m$.

(iii) For every $B \in R^m$, there exists $X \in \mathbb{R}^n$ such that $AX = B$.

(iv) For every $B \in R^m$, the system $AX = B$ is consistent.

(v) The rank of A equals m.

(vi) There is a leading 1 in every row of Q.

Proof: It is clear that (i), (ii), (iii) and (iv) are all different ways of saying the same thing, namely that for any vector B in $\mathbb{R}^m$ there is a list of coefficients X so that
$B \ = \ x_1 C_1 + x_2 C_2 + \ldots x_n C_n.$

Since there is at most one leading 1 in any row of Q, it is clear that (v) and (vi) are saying the same thing.

We can write $Q = UA$ where U is an invertible $n \times n$ matrix, the product of the elementary matrices for the operations which move A to Q.

Now assume that there is a leading 1 in every row of Q. For any B, we put the augmented matrix $(A|B)$ in Reduced Echelon Form by multiplying by U to get $(Q|UB)$. Since there is a leading 1 in every row of Q there is no leading 1 in the last column and the system is consistent. That is, (vi) implies (iv).

On the other hand, if Q has a row of zeroes, then $(Q|\mathbf{e}_m)$ is inconsistent and so $(A|U^{-1}\mathbf{e}_m)$ is inconsistent. So not (vi)

Let us look at Exercises 5.1/ 2a, 3a,   page 267

Linear Independence

A list $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *is* (not *are*) *linearly independent* (= li) if $c_1\mathbf{v}_1 + ... + c_n\mathbf{v}_n \ = \ \mathbf{0}$ only if $c_1 = c_2 ... = c_n = 0$. That is, the only way to get $\mathbf{0}$ as a linear combination is by using the zero coefficients.

By the Fundamental Principle: The list S is li if and only if we can *equate coefficients*. That is:

$$\begin{aligned} c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \ &= \ a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \\ \Longrightarrow \qquad & c_1 = a_1, \ldots, c_n = a_n. \end{aligned} \tag{4.10}$$

By convention, we regard the empty list as linearly independent.

It is important to think of $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ as a list and not a set. The order in which we write the vectors does not matter, but vectors may be repeated. For example, the list consisting of a single vector $\{\mathbf{w}\}$ is li if and only if $\mathbf{w} \neq \mathbf{0}$. (Why?)

On the other hand, If in the list, $\mathbf{v}_1 = \mathbf{v}_2$ then the list is not li because $1\mathbf{v}_1 \ + \ (-1)\mathbf{v}_2 \ + \ 0\mathbf{v}_3 + \cdots + 0\mathbf{v}_n = \mathbf{0}$.

Notice that if a set spans, then any larger set spans, while if a list is li then any list contained in it is li.

That is, for $S \subset T$:

- If S spans, then T spans.
- If T is li, then S is li.

When $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is not linearly independent, it is called *linearly dependent* (= ld). Thus, S is ld when the zero vector can be written as a non-trivial linear combination. There exists a list of coefficients not all zero such that
$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \;=\; \mathbf{0}$.

The list S is ld if one vector on the list is a linear combination of the others. For example, if $\mathbf{v}_n = c_1\mathbf{v}_1 + \cdots + c_{n-1}\mathbf{v}_{n-1}$ then $\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_{n-1}\mathbf{v}_{n-1} - \mathbf{v}_n$ and one of the coefficients is certainly not zero. (Which one and what is it?)

Conversely,

Theorem 4.04: For the lists $S \;=\; \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $T = \{\mathbf{v}_{n+1}, \ldots, \mathbf{v}_m\}$, assume that S is li and $S \cup T \;=\; \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_m\}$ ld. There exists a vector $\mathbf{v}_i$ in T which is a linear combination of the remaining vectors of $S \cup T$. In addition, $span(S \cup T) = span(S \cup T \setminus \{\mathbf{v}_i\})$.

Proof: Assume $c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$ and not all the c_i's equal 0.

We can *solve for* $\mathbf{v}_i$ if c_i is not zero. For example, if $c_m \neq 0$, then $\mathbf{v}_m = -(1/c_m)(c_1\mathbf{v}_1 + \cdots + c_{m-1}\mathbf{v}_{m-1})$.

If all of $c_{n+1} = \cdots = c_m = 0$, then $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$ would imply $c_1 = \cdots = c_n = 0$ because S is li.

So one of the c_i's with $n < i \leq m$ is nonzero and we can solve for that $\mathbf{v}_i$.

Thus $\{\mathbf{v}_i\} \subset span(S \cup T \setminus \{\mathbf{v}_i\})$ and so $span(S \cup T) = span(S \cup T \setminus \{\mathbf{v}_i\})$ by (iii) of the Corollary 02.
$\square$

Theorem 4.05: Assume that the list $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is li and $T = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ spans V. We can renumber T so that $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_{n+1}, \ldots, \mathbf{w}_m\}$ spans V. That is, we can replace n of the vectors in T by the vectors of S and still have a spanning set. In particular, $m \geq n$.

Proof Suppose that $B_k = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{w}_{k+1}, \ldots, \mathbf{w}_m\}$ spans V for some k with $0 \leq k < n$. Here B_0 is just T.

Because B_k spans, $\mathbf{v}_{k+1}$ is a linear combination of the vectors of B_k and so

$$B_k' = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \mathbf{w}_{k+1}, \ldots, \mathbf{w}_m\} = B_k \cup \{\mathbf{v}_{k+1}\}$$

is linearly dependent and of course it still spans.

Now we apply Theorem 04 which implies that one of the vectors $\mathbf{w}_{k+1}, \ldots, \mathbf{w}_m$ is a linear combination of the vectors B'_k. By renumbering, we can assume that it is $\mathbf{w}_{k+1}$ which lies in $span(B_{k+1})$ with

$$B_{k+1} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \mathbf{w}_{k+2}, \ldots, \mathbf{w}_m\} = B'_k \setminus \{\mathbf{w}_{k+1}\}.$$

By Theorem 04 $span(B_{k+1}) = span(B'_k) = V$.

Repeat the process until we get to $B_n = B$. Technically, this is a proof by induction.
$\square$

Theorem 4.06: Let A be an $m \times n$ matrix with columns $C_1, \ldots C_n \in \mathbb{R}^m$. Assume that Q in Reduced Echelon Form is row equivalent to A. The following conditions are equivalent.

(i) The list $\{C_1, \ldots C_n\}$ of the columns of A is linearly independent.

(ii) The homogeneous system $AX = \mathbf{0}$ has only the trivial solution $X = \mathbf{0}$.

(iii) For every $B \in R^m$, there exists at most one $X \in \mathbb{R}^n$ such that $AX = B$.

(iv) The rank r of A equals n.

(v) There is a leading 1 in every column of Q.

Proof: (i) and (ii) are saying the same thing, and (ii) says that for $B = \mathbf{0}$ the system has only one solution. Hence, (iii) implies (ii). On the other hand if $AX_1 = B = AX_2$ with $X_1 \neq X_2$, then $X = X_1 - X_2$ is a nontrivial solution of $AX = \mathbf{0}$.

Since there is at most one leading 1 in every column, (iv) and (v) are saying the same thing.

By the Gaussian algorithm, the solutions of $AX = \mathbf{0}$ are the same as those of $QX = \mathbf{0}$.

The solution $X = \mathbf{0}$ is unique when every variable is a fixed variable, that is, $n = r$.

If $r < n$, then there are free variables and $AX = \mathbf{0}$ has infinitely many solutions.

BE ABLE TO DEFINE: $S = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is li or ld.

If in the list $S = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ some $\mathbf{v}_i = \mathbf{0}$, then the list is ld. (Why?)

If $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ then the pair is ld if and only if one of the two vectors is a scalar multiple of the other. (Why?)

Let us look at Exercises 5.2/1a,8 pages 278, 279, and 6.3/ 1ac, page 349.

Basis and Dimension

A list B of vectors in V is a *basis* if it spans and is li.

We call a vector space *finite dimensional* when it admits a finite spanning set.

Theorem 5.01 Dimension Theorem: If the list $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is li in V and $T = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ spans V, then $m \geq n$. In particular, if V is finite dimensional, then any li list is finite.

If both S and T are bases, i.e. both are li lists which span V, then $m = n$.

Proof: We saw in Theorem 4.05 that $m \geq n$.

If V has a spanning set of size m, then since every li list has size at most m, there can't be an infinite li list,

If both are bases then by Theorem 4.05 again $m \geq n$ and $n \geq m$.

$\square$

For a finite dimensional vector space, the number of vectors in any basis is called its *dimension*. For a finite set S we will write $\#S$ for the number of elements in it. So V has dimension n when $\#B = n$ for any basis B of V.

The empty set is li by convention and $span(\emptyset) = \{\mathbf{0}\}$. Hence, $\emptyset$ is a basis for $\{\mathbf{0}\}$. and so the dimension of $\{\mathbf{0}\}$ is 0.

If $\mathbf{v}$ is a nonzero vector in V, then $\mathbb{R}v \; = \; \{c\mathbf{v} : c \in \mathbb{R}\}$ is a subspace of V with basis $\{v\}$. Hence, $\mathbb{R}v$ has dimension 1.

Theorem 5.02 (*Between Theorem*): For the lists $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $T = \{\mathbf{v}_{n+1}, \ldots, \mathbf{v}_m\}$, assume that S is li and $S \cup T = \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_m\}$ spans.

There exists a basis B with $S \subset B \subset S \cup T$. That is, the basis consists of all of the vectors of S together with some (maybe none, maybe all) of the vectors of T.

Proof: Look first, at two extreme cases.

If S spans then since it is li, it is a basis. Use $B = S$, using none of the vectors of T. For example, if $m = 0$ so that there are no vectors in T, then $S = S \cup T$ spans and $B = S$ is a basis.

If $S \cup T$ is li then since it spans, it is a basis. Use $B = S \cup T$, using all of the vectors of T.

Now if $S \cup T$ is not li, then by Theeorem 04 one of the vectors of T can be written as a linear combination of the remaining vectors of $S \cup T$. By renumbering we can suppose this is true of the last vector $\mathbf{w}_m$. Call $T' = = \{\mathbf{w}_1, ..., \mathbf{w}_{m-1}\}$. So that last vector $\mathbf{w}_m \in span(S \cup T')$. As mentioned in Theorem 04, (iii) of the Corollary 02 implies that $span(S \cup T') = span(S \cup T)$. But $S \cup T$ spans V and so $S \cup T'$ spans V.

We continue, doing a loop. If $S \cup T'$ is li then we are done with $B = S \cup T'$. If $S \cup T'$ is ld we can remove some vector from T' and still have a spanning set.

Eventually, we get to a situation where together the set of remaining vectors is li and so we have the basis B. If not before this will happen when we have removed all of the vectors in T.

Technically, we are using mathematical induction on m. Remember that for the initial case $m = 0$ is easy. In that case, $B = S$ is a basis.

Assume the result for $m = k$ and suppose that $m = k + 1$.

We saw that if $S \cup T$ is li, then $B = S \cup T$ is a basis, and if $S \cup T$ is ld, we can remove a vector and still have the spanning set $S \cup T'$ but now with m for T' being k. The induction hypothesis implies that there is a basis B with $S \subset B \subset S \cup T' \subset S \cup T$.

Result then follows from the Principle of Mathematical Induction.
$\square$

Corollary 5.03: Any finite spanning set for a vector space contains a basis.

Any linearly independent list for a finite dimensional vector space extends to a basis.

Proof: If T spans apply Theorem 08 with $S = \emptyset$.

If S is a li list and T is any finite spanning set, then S too is finite by the Dimension Theorem. Apply Theorem 08 with S and T.

$\square$

Corollary 5.04: If W is a proper subspace of a finite dimensional space V, i.e. $W \neq V$, then $dimW < dimV$.

Proof: If B is a basis for W, then B is an li list of vectors in V which does not span V because it spans W. By Corollary 5.03 we can extend B to a basis B' of V and B' contains more vectors than just those in B. So

$dimW = \#B < \#B' = dimV$.

Theorem 5.05: Let S be a list of vectors in a vector space V of dimension n.

(a) If S is li then $\#S \leq n$ with equality if and only if S is a basis.

(b) If S spans then $\#S \geq n$ with equality if and only if S is a basis.

(c) Any two of the following conditions implies the third.

   (i) $\#S \quad = \quad n.$

   (ii) S is li.

   (iii) S spans.

Proof:(a) Use Corollary 5.03 to extend S to a basis B.
$\#S \leq \#B = n$.

(b) Use Corollary 5.03 to shrink S to a basis B.
$\#S \geq \#B = n$.

In either case, if $\#S = n = \#B$ then $S = B$ and so S is a basis.

(c) If S is li and spans then it is a basis and so has size n. For the other directions apply (a) and (b).
$\square$

Coordinates

We began our serious study of linear algebra by introducing a new meaning for the term *vector*. Instead of something with magnitude and direction, a vector is an element of a vector space.

Now we alter the meaning of the term *coordinates*.

If you were asked: What are the coordinates of the vector $(3, 0, -1, 5) \in \mathbb{R}^4$? You would have given the obvious answer: $3, 0, -1, 5$.

Now the correct response, is to ask another question: Coordinates with respect to what basis?

If $B = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ is a basis for V (so that V has dimension n) and $\mathbf{v}$ is a vector in V, then there is a unique list of numbers $c_1, ..., c_n$ such that $\mathbf{v} = c_1\mathbf{v}_1 + ... + c_n\mathbf{v}_n$.

[How do you know that c_i's exist? How do you know they are unique?]

The numbers $\{c_1, ..., c_n\}$ are the *coordinates of* **v** *with respect to the basis* B.

The vector

$$[\mathbf{v}]_B = \begin{pmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_n \end{pmatrix} \in \mathbb{R}^n \tag{5.1}$$

is <u>the</u> *coordinate vector of* **v** *with respect to* B.

In many cases, a vector space is described using a linear list of parameters. For example,

$$M_{22} \ = \ \{A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \ a, b, c, d \in \mathbb{R} \ \}$$

Here there are four parameters, four free variables each of which can take on an arbitrary real value independent of the choice for the others. We think of the coordinates of A as being a, b, c, d. In such a case there is an associated *standard basis*. Its elements are each obtained by setting one the parameters equal to 1 and the others equal to 0.

$$\begin{aligned} S \ = \ \{E_a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ E_b = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ E_c = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_d = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \ \}. \end{aligned} \tag{5.2}$$

Clearly, $A = aE_a + bE_b + cE_c + dE_d$.

Since two matrices are equal if and only if the corresponding entries are equal, we can equate coefficients and so S is an li list. It clearly spans and so is a basis for M_{22}

From (5.2) we have

$$[A]_S = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

If you choose a different order for the parameters, e.g. a, c, b, d, then the ordering of the list in S and so the order of the entries in $[A]_S$ would be changed. Any choice of order is ok, but having chosen one, it should be kept consistently.

In general, the space M_{mn} of $m \times n$ matrices has dimension mn which is the number of entries and so is the number of elements of the standard basis.

For $\mathbb{R}^n$ itself with a typical vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

the standard basis is

$$S = \{\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 1 \end{pmatrix}\} \tag{5.3}$$

and $\mathbf{x}$ is its own coordinate vector with respect to the standard basis S.

The dimension of $\mathbb{R}^n$ is n.

Let $\mathcal{P}_n$ be the space polynomials of degree at most n. So a typical polynomial is given by $p(t) \;= a_0 + a_1 t + ... + a_n t^n$. The parameters are the coefficients $(a_0, ..., a_n)$.

If $p(t) = q(t) = b_0 + b_1 t + ... + b_n t^n$, then we can equate coefficients to get $a_0 = b_0, \ldots, a_n = b_n$.

To see this, set $t = 0$ to get $a_0 = b_0$, then takes the derivative and set $t = 0$ to get $a_1 = b_1$, and continue until at the n^{th} derivative we get $n!a_n = n!b_n$.

Treating the coefficients as parameters we get the standard basis $S = \{1, t, ..., t^n\}$ and

$$[p(t)]_S \ = \ \begin{pmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$$

There are $n+1$ vectors in this standard basis and so the dimension of $\mathcal{P}_n$ is $n+1$.

For a fixed $m \times n$ matrix, the solution space of the homogeneous system $AX = \mathbf{0}$ is a subset of $\mathbb{R}^n$ called the nullspace of A, denoted $Null(A)$. Observe that $AX_1 = \mathbf{0}$ and $AX_2 = \mathbf{0}$ implies $A(X_1 + X_2) = \mathbf{0}$ and $A(cX_1) = \mathbf{0}$. Thus, $Null(A)$ is a subspace of $\mathbb{R}^n$. Notice that if $B \neq \mathbf{0}$, then the solution set of $AX = B$ in $\mathbb{R}^n$ is <u>not</u> a subspace.

Recall how we solve the system $AX = B$.

Let Q be the matrix in Reduced Row Echelon Form which is row equivalent to A, so that $Q = UA$ with U the invertible $m \times m$ matrix which is the product of the elementary matrices for the row operations that we used to get from A to Q.

The vector Y is in the column space of A, i.e. it is a linear combination of the columns, if and only if the system $AX = Y$ is consistent. This is the same as when $QX = UY$ is consistent because these two systems have the same solutions. This in turn means that any row of zeroes for Q extends to a 0 in the corresponding row of UY.

Thus, the augmented matric $(Q|UY)$ is in Reduced Echelon Form and all of the leading 1's occur in the columns of Q. There are r of these where r is the rank of A. These r columns correspond to the fixed variables or pivot variables The remaining $n - r$ columns of Q correspond to free variables. That is, each is a parameter whose value can be chosen arbitrarily and independent of the others.

Theorem 5.06: Let A be an $m \times n$ matrix of rank r, with $Q = UA$ the matrix row equivalent to A which is in Reduced Echelon Form using the invertible $m \times m$ matrix U.

Let $B = \{C_{i_1}(A), C_{i_2}(A), \dots, C_{i_r}(A)\}$ be the list of the columns of A which correspond to the fixed variables. That is, these are the columns of A in which the leading 1 occur in Q after the row operations. The list B is a basis for the column space $Col(A)$. Thus, the dimension of $Col(A)$ is the rank r.

For the nullspace of A the free variables are parameters for the general solution of $AX = \mathbf{0}$. The $n - r$ parameters provide a standard basis for $Null(A)$ and so the dimension of $Null(A)$ is $n - r$.

Proof: For any $Y \in Col(A)$ there is a unique solution of the system $AX = Y$ with all of the parameters chosen to be 0. That is, every coefficient except those on $C_{i_1}(A), C_{i_2}(A), \ldots, C_{i_r}(A)$ is 0. This yields any Y as a linear combination of the columns in B. That is, B spans the column space. Uniqueness implies that we can equate coefficients when comparing linear combinations of the columns in B. This in turn means that B is an li list. Since B is a basis for $Col(A)$, the dimension equals r.

In particular, the trivial solution with all x_i's equal to 0 is the only solution of $AX = \mathbf{0}$ with all the parameters chosen to be 0.

The list of parameters is the list of coordinates of an element of $Null(A)$ with respect to the associated standard basis. There are $n - r$ parameters and so there are $n - r$ elements of the associated basis.

$\square$

Consider the example:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

The column space is $\mathbb{R}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ with basis $\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$.

The general solution of $AX = \mathbf{0}$ is given by $x_3 = r, x_2 = s, x_1 = -2s - 3r$. So the basis of the nullspace is

$$\{X_s = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \ X_r = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}\}$$

Notice that $Col(A)$ is a different subspace from $Col(Q)$ although they both have the same dimension, namely their common rank.

We have just seen that even if A is row equivalent to Q, the column spaces may differ. They usually do. On the other hand, the row space of A, the span of the rows of A in $\mathbb{R}^n$, denoted $Row(A)$, is always the same as $Row(Q)$.

Observe that if you rearrange the order of the list of vectors, the span, the set of linear combinations is unchanged. The same is true if you replace a vector on the list by a nonzero multiple of it. Finally, if you replace $\mathbf{v}_i$ by $\mathbf{v}_i + c\mathbf{v}_j$ for $j \neq i$, the span does not change. For example,

$$\begin{aligned} c_1(\mathbf{v}_1 + c\mathbf{v}_2) + c_2\mathbf{v}_2 + \ldots c_n\mathbf{v}_n &= c_1\mathbf{v}_1 + (c_1c + c_2)\mathbf{v}_2 + \ldots c_n\mathbf{v}_n, \\ d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \ldots c_n\mathbf{v}_n &= d_1(\mathbf{v}_1 + c\mathbf{v}_2) + (d_2 - d_1c)\mathbf{v}_2 + \ldots c_n\mathbf{v}_n. \end{aligned}$$

Theorem 5.07: Let A be an $m \times n$ matrix of rank r, with Q the matrix row equivalent to A which is in Reduced Echelon Form. The row spaces $Row(A)$ and $Row(Q)$ are the same. The list of nonzero rows in Q is a basis for this common row space. As there are r such rows, the dimension of $Row(A)$ is the rank r.

Proof: We have just seen that the row operations do not change the row space and so $Row(A) = Row(Q)$. The nonzero rows $\{R_1(Q), R_2(Q), \ldots, R_r(Q)\}$ clearly span the row space.

Suppose the leading 1 occur in columns $i_1, \ldots, i_r$ and so in position $(1, i_1), (2, i_2), \ldots, (r, i_r)$. Recall that the leading 1 is the only nonzero entry in its column since Q is in Reduced Echelon Form. It therefore follows that the linear combination $Z = c_1 R_1 + c_2 R_2 + \ldots c_r R_r$ equals c_ℓ in the ℓ place. Hence, $Z = \mathbf{0}$ implies all the c_ℓ's equal 0. That is, $\{R_1, \ldots, R_r\}$ is an li list and so is a basis for $Row(A) = Row(Q)$.
$\square$

Theorem 5.08: For A be an $n \times n$ matrix, the following are equivalent.

(i) A is invertible.

(ii) The columns of A form an li list.

(iii) The dimension of the column space $Col(A)$ is n.

(iv) The rows of A form an li list.

(v) The dimension of the row space $Row(A)$ is n.

(vi) The null space $Null(A)$ equals $\{\mathbf{0}\}$.

Proof: From the Fredholm Alternative, A is invertible if and only if the rank $r = n$.

By Theorem 5.06, the number of li columns equals the rank r and this is the dimension of $Col(A)$. So both (ii) and (iii) are equivalent to $r = n$.

By Theorem 5.07, the number of li rows equals the rank r and this is the dimension of $Row(A)$. So both (iv) and (v) are equivalent to $r = n$.

By Theorem 5.06 again the dimension of $Null(A)$ is $n - r$. $Null(A) = \{\mathbf{0}\}$ if and only if $dimNull(A) = 0$ and so if and only if $r = n$.
$\square$

BE ABLE TO DEFINE:(1) B is a basis for V.
(2) The dimension of V.
(3) The coordinates of a vector $\mathbf{v}$ of V with respect to a basis for V.

For an $m \times n$ matrix A be able to get a basis for $Col(A), Null(A)$ and $Row(A)$.

Let us look at
Exercises 6.3/ 5b, 6c, 10b, 24, pages 350, 351.
6.4/ 2b, 3a pages 358, 359.

Linear Transformations

Earlier we define a linear transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by multiplying the $m \times n$ matrix A.

In general, a *linear transformation* or *linear map* is a function $T : V \to W$ between the vector spaces V and W. That is, the inputs and outputs are vectors, and T satisfies *linearity*, which is also called the *superposition property*:

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) \quad \text{and} \quad T(a\mathbf{v}) = aT(\mathbf{v}). \tag{6.1}$$

In particular, $T(\mathbf{0}) = T(0\mathbf{0}) = 0T(\mathbf{0}) = \mathbf{0}$, and $T(-\mathbf{v}) = T((-1)\mathbf{v}) = (-1)T(\mathbf{v}) = -T(\mathbf{v})$.
It follows that T relates linear combinations:

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots\ c_k\mathbf{v}_k) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \ldots c_kL(\mathbf{v}_k). \tag{6.2}$$

This property of linearity is very special. It is a standard algebra mistake to apply it to functions like the square root function and sin and cos etc. for which it does not hold. On the other hand, these should be familiar properties from calculus. The operator D associating to a differentiable function f its derivative Df is a most important example of a linear operator.

From a linear map we get an important examples of subspaces.

For a linear map $T : V \to W$, the set of vectors $\{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}$ *solution space of the homogeneous equation*, is a subspace of V called the kernel of T, $Ker(T)$. If $\mathbf{r}$ is not $\mathbf{0}$ then the solution space of $T(\mathbf{v}) = \mathbf{r}$ is not a subspace. For example, it does not contain $\mathbf{0}$.

For a linear map $T : V \to W$, the set of vectors $\{\mathbf{w} \in W :$ for some $\mathbf{v} \in V, \ T(\mathbf{v}) \ = \mathbf{w}\}$ is called *the image of L*, denoted $Im(T)$.

Check that $Ker(T) \subset V$ and $Im(T) \subset W$ are subspaces.

For the linear map $T_A : \mathbb{R}^n \to \mathbb{R}^m$ associated with the $m \times n$ matrix A, $Ker(T_A) = Null(A)$ and $Im(T_A) = Col(A)$.

Theorem 6.01: (a) If $T : V \to W$ and $S : W \to U$ are linear maps, then the composition $S \circ T : V \to U$ defined by $S \circ T(\mathbf{v}) = S(T(\mathbf{v}))$ is a linear map.

(b) A linear map $T : V \to W$ is one-to-one if and only if $Ker(T) = \{\mathbf{0}\}$.

(c) A linear map $T : V \to W$ is onto if and only if $Im(T) = W$.

(d) If a linear map $T : V \to W$ is one-to-one and onto, then the inverse map $T^{-1} : W \to V$ defined by

$$T^{-1}(\mathbf{w}) = \mathbf{v} \quad \Leftrightarrow \quad T(\mathbf{v}) = \mathbf{w} \tag{6.3}$$

is a linear map.

A one-to-one, onto linear map is called a *linear isomorphism*.

Proof: (a)
$S(T(c\mathbf{v}_1+\mathbf{v}_2) = S(cT(\mathbf{v}_1)+T(\mathbf{v}_2)) = cS(T(\mathbf{v}_1))+S(T(\mathbf{v}_2)).$

(b) $T(\mathbf{0}) = \mathbf{0}$ and so if T is one-to-one, $Ker(T) = \{\mathbf{0}\}$. Conversely, if $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, then $T(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}$. So if $Ker(T) = \{\mathbf{0}\}$, we have $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$ and so $\mathbf{v}_1 = \mathbf{v}_2$.

(c) This is clear from the definition of $Im(T)$.

(d) $T^{-1}(c\mathbf{w}_1+\mathbf{w}_2) = c\mathbf{v}_1+\mathbf{v}_2 \quad \Leftrightarrow \quad c\mathbf{w}_1+\mathbf{w}_2 = T(c\mathbf{v}_1+\mathbf{v}_2)$.
$\square$

Theorem 6.02: Let $T : V \to W$ be a linear map and $D = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a list of vectors in V. Define $T(D) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$

(a) If $T(D)$ is an li list in W, then D is an li list in V.

(b) If D is an li list in V and $Ker(T) = \{\mathbf{0}\}$, then If $T(D)$ is an li list in W.

(c) If D spans V and $Im(T) = W$, then $T(D)$ spans W.

(d) If D is a basis for V and T is a linear isomorphism, then $T(D)$ is a basis for W.

(e) If D spans V and $S : V \to W$ is a linear map with $T(\mathbf{v}_1) = S(\mathbf{v}_2), \ldots, T(\mathbf{v}_1) = S(\mathbf{v}_2)$, then $T = S$. That is, $T(\mathbf{v}) = S(\mathbf{v})$ for all $\mathbf{v} \in V$.

Proof: (a) Assume $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$. We must show $c_1 = \cdots = c_n = 0$.

$\mathbf{0} = T(\mathbf{0}) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1 T(\mathbf{v}_1) + \cdots + c_n T(\mathbf{v}_n)$. Because the list $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is li, it follows that $c_1 = \cdots = c_n = 0$.

(b) Assume $\mathbf{0} = c_1 T(\mathbf{v}_1) + \cdots + c_n T(\mathbf{v}_n) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n)$. Because $Ker(T) = \{\mathbf{0}\}$, $\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. Because D is li, $c_1 = \cdots = c_n = 0$.

(c) For $\mathbf{w} \in W$, there exists $\mathbf{v}$ with $T(\mathbf{v}) = \mathbf{w}$ (Why?). There exist coefficients so that $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$ (Why?). So $\mathbf{w} = c_1 T(\mathbf{v}_1) + \cdots + c_n T(\mathbf{v}_n)$.

(d) follows from (b) and (c).

(e) Check that $\{\mathbf{v} : T(\mathbf{v}) = S(\mathbf{v})\}$ is a subspace of V because S and T are linear. The subspace contains the spanning set D and so equals all of V.
□

Theorem 6.03: For $T : V \to W$ a linear map,

$$dimV \;=\; dimKer(T) \;+\; dimIm(T). \tag{6.4}$$

Proof: Every vector of $Im(T)$ is of the form $T(\mathbf{v})$ and so we can choose a basis $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_r)\}$ for $Im(T)$. Let $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ be a basis for $Ker(T)$. We will show that $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots \mathbf{v}_n\}$ is a basis for V which will show that $n = dimV$.

Assume $\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r + c_{r+1}\mathbf{v}_{r+1} + \ldots c_n\mathbf{v}_n$. We must show $c_1 = \ldots c_r = c_{r+1} \cdots = c_n = 0$.

Apply T and note the $\mathbf{v}_i \in Ker(T)$ for $r < i \leq n$ implies $\mathbf{0} = c_1 T(\mathbf{v}_1) + \cdots + c_r T(\mathbf{v}_r)$. So $c_1 = \cdots = c_r = 0$ (Why?)

This implies that $\mathbf{0} = c_{r+1}\mathbf{v}_{r+1} + \ldots c_n\mathbf{v}_n$. So $c_{r+1} = \cdots = c_n = 0$ (Why?)

Thus, $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots \mathbf{v}_n\}$ is li.

Now let $\mathbf{v} \in V$. We must find coefficients so that $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r + c_{r+1}\mathbf{v}_{r+1} + \ldots c_n\mathbf{v}_n$.

There exist coefficients so that $T(\mathbf{v}) = c_1 T(\mathbf{v}_1) + \cdots + c_r T(\mathbf{v}_r)$ (Why?)

$$T(\mathbf{v} - c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r) = T(\mathbf{v}) - (c_1 T(\mathbf{v}_1) + \cdots + c_r T(\mathbf{v}_r)) = \mathbf{0}$$

and so there exist coefficients such that

$$\mathbf{v} - c_1\mathbf{v}_1 + \cdots + c_r\mathbf{v}_r = c_{r+1}\mathbf{v}_{r+1} + \ldots c_n\mathbf{v}_n.$$

Thus, $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots \mathbf{v}_n\}$ spans V.

$\square$

Theorem 6.04: Let $D = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a list of vectors in V. The map $T_D : \mathbb{R}^n \to V$ defined by

$$T_D(x_1, \ldots, x_n) \;=\; \sum_{i=1}^{n} x_1\mathbf{v}_i \;=\; x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n \qquad (6.5)$$

is a linear map.

If D is a basis, then T_D is a linear isomorphism and inverse map $T_D^{-1} : V \to \mathbb{R}^n$ is given by $T_D^{-1}(v) \;=\; [\mathbf{v}]_D$, the coordinate vector of $\mathbf{v}$ with respect to the basis D.

Proof: We saw above that the sum of two linear combinations on a list is the linear combination obtained by adding the corresponding coefficients. Similarly,

$T_D(cx_1, \ldots, cx_n) = cT_D(x_1, \ldots, x_n)$. Thus, T_D is linear.

If D is a basis then for any $\mathbf{v} \in V$ the equation $\mathbf{v} \;=\; x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$ uniquely defines the coefficients and the list of coefficients is the coordinate vector $[\mathbf{v}]_D$.
$\square$

Corollary 6.05: For finite dimensional vector spaces V and W, there is a linear isomorphism $T : V \to W$ if and only if $dimV = dimW$. In particular, if $dimV = n$, then there is a linear isomorphism $T : \mathbb{R}^n \to V$.

Proof: If $D = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V, then $T_D : \mathbb{R}^n \to V$ is a linear isomorphism and such a basis exists exactly when $dimV = n$.

Thus, if $dimV = dimW = n$ then there exist linear isomorphisms $T : V \to \mathbb{R}^n$ and $S : W \to \mathbb{R}^n$ and so the composition $S^{-1} \circ T : V \to W$ is a linear isomorphism.

On the other hand, if $T : V \to W$ is a linear isomorphism and D is a basis for V, then by Theorem 6.02(d) $T(D)$ is a basis for W. Therefore, $dimV = \#D = \#T(D) = dimW$.
□

Let us look at Exercise 7.2/ 1b, page 385.

Matrix of a Linear Transformation

Recall that for A an $m \times n$ matrix the linear map $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is defined by $T_A(X) = AX$. By using bases we can represent every linear map between finite dimensional vector spaces in this way.

If $T : V \to W$ is a linear map and $B = \{\mathbf{v}_1, ..., \mathbf{v}_n\}, D = \{\mathbf{w}_1, ..., \mathbf{w}_m\}$ are bases for V and W, respectively, then the matrix $[T]_{DB}$ is the $m \times n$ matrix given by

$$[T]_{DB} = [\ [T(\mathbf{v}_1)]_D ... [T(\mathbf{v}_n)]_D\ ]. \tag{6.6}$$

That is, we form the matrix by applying T to each of the domain basis vectors from B in V. We list them in order, thinking of them as a matrix but with vectors in W instead of columns of numbers. We convert each vector to an actual column of numbers by replacing each by its column of D coordinates. Thus, we obtain the $m \times n$ matrix $[T]_{DB}$.

Theorem 6.06: Let $T : V \to W$ be a linear map with $B = \{\mathbf{v}_1, ..., \mathbf{v}_n\}, D = \{\mathbf{w}_1, ..., \mathbf{w}_m\}$ bases for V and W. Let $[T]_{DB}$ be the $m \times n$ matrix associated to the linear map by the choice of bases. If $\mathbf{v} \in V$, then

$$[T(\mathbf{v})]_D = [T]_{DB}[\mathbf{v}]_B. \tag{6.7}$$

That is, the D coordinate vector of $\mathbf{w} = T(\mathbf{v})$ in $\mathbb{R}^m$ is

obtained by multiplying the B coordinate vector of $\mathbf{v}$ in $\mathbb{R}^n$ by the $m \times n$ matrix $[T]_{DB}$.

Proof: By definition $[\mathbf{v}]_B = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$ means $\mathbf{v} = x_1\mathbf{v}_1 + \ldots x_n\mathbf{v}_n$,

and so $\mathbf{w} = T(\mathbf{v}) = x_1 T(\mathbf{v}_1) + \ldots x_n T(\mathbf{v}_n)$. By Theorem 6.04, the coordinate map $\mathbf{w} \mapsto [\mathbf{w}]_D$ is linear and so

$$[\mathbf{w}]_D \;=\; x_1[T(\mathbf{v}_1)]_D + \ldots x_n[T(\mathbf{v}_n)]_D.$$

This is, the linear combination of the columns of $[T]_{DB}$ with coefficients $x_1, \ldots, x_n$. That is exactly

$$[T]_{DB} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} \;=\; [T]_{DB}[\mathbf{v}]_B.$$

$\square$

Corollary 6.07: Let $T : V \to W$ and $S : W \to U$ be linear maps with B, D, E bases for V, W and U.

$$[S \circ T]_{EB} \ = \ [S]_{ED}[T]_{DB}. \tag{6.8}$$

That is, the matrix of the composed linear map $S \circ T$ is the

product of the matrices of S and T provided that the same basis D is used for W as the range of T and as the domain of S.

Proof: Let $\mathbf{v}$ be an arbitrary vector in V. By Theorem 6.06 applied first to $S \circ T$, then to S and then to T we have

$$\begin{aligned} [S \circ T]_{EB}[\mathbf{v}]_B \ = \ [(S \circ T)(\mathbf{v})]_E \ = \\ [(S(T(\mathbf{v}))]_E \ = \ [S]_{ED}[T(\mathbf{v})]_D \ = [S]_{ED}[T]_{DB}[\mathbf{v}]_B. \end{aligned} \tag{6.9}$$

The result follows because if A and B are $m \times n$ matrices such that $AX = BX$ for all $X \in \mathbb{R}^n$, then $A = B$. (Hint: let X vary over the columns of I_n, which list is the standard basis for $\mathbb{R}^n$).

$\square$

An important special case lets us change the coordinates from one basis to another. We use the identity map I on the vector space V, so that $I(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v}$ in V, but we use different bases on the domain and range.
Let $B = \{\mathbf{v}_1, ..., \mathbf{v}_n\}, D = \{\mathbf{w}_1, ..., \mathbf{w}_n\}$ be two different bases on a vector space V. They have the same number n of elements when $dim V = n$. The transition matrix from B to D is given by:

$$[I]_{DB} = [\ [\mathbf{v}_1]_D ... [\mathbf{v}_n)]_D\ ]. \tag{6.10}$$

That is, the columns are the D coordinates of the B vectors listed in order.

Corollary 6.08: Let B and D be bases for a vector space V of dimension n.

(a) $[I]_{BB} = I_n$. That is, the transition matrix from a basis to itself is the identity matrix.

(b) $[I]_{BD} = ([I]_{DB})^{-1}$. That is, the transition matrix from D to B is the inverse matrix of the transition matrix from B to D.

(c) For any vector $\mathbf{v} \in V$,

$$[\mathbf{v}]_D = [I]_{DB}[\mathbf{v}]_B. \tag{6.11}$$

Proof: (a) is easy to check, e.g. $\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n$. Then (b) follows from Corollary 6.07.

Finally, (c) is a special case of Theorem 6.06.
□

As we have seen, many of the spaces we look at have a standard basis S whose coordinate vectors are easy to read off. If $T : V \to W$ is a linear map with B is a basis for V and S is a standard basis for W, then it is easy to compute $[T]_{SB}$.

For example, if A is an $m \times n$ matrix and $X \in \mathbb{R}^n$, then with respect to the standard bases S_n on $\mathbb{R}^n$ and S_m on R^m, just as the coordinate vector $[X]_{S_n}$ is X itself, so too $[T_A]_{S_m S_n} = A$.

If $T : V \to W$ is a linear map with $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}, D = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ bases for V and W and S is a standard basis for W, then usually $[T]_{SB}$ and $[I]_{SD}$ are easy to read directly. It is then sometimes easiest to use the following application of Corollaries 6.07 and 6.08:

$$[T]_{DB} = [I]_{DS}[T]_{SB} = ([I]_{SD})^{-1}[T]_{SB}. \tag{6.12}$$

Let us look at Exercises 9.1/1ad, page 501.