Notes 14 – Bases and Linear Mappings

For the rest of the course, students should have in mind the following vector spaces: $\mathbb{R}^n$ (that is, either $\mathbb{R}^{1,n}$ or $\mathbb{R}^{n,1}$), $\mathbb{R}^{m,n}$ and $\mathbb{R}_n[x]$. All are real vector spaces, with $F = \mathbb{R}$. Many other vector spaces can then be defined by choosing subspaces, a concept that we have already investigated in $\mathbb{R}^n$.

L14.1 Linear combinations and subspaces. Let $\mathbf{u}_1, \cdots, \mathbf{u}_k$ be elements of a vector space V. We can use the same notation as before for the set of all *linear combinations* (LC's) of the vectors listed, so that

$$\mathscr{L}\{\mathbf{u}_1, \cdots, \mathbf{u}_k\} = \{a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k \ : \ a_i \in F\}. \tag{1}$$

The only novelty is that the coefficients now belong to F (if this is different from $\mathbb{R}$). Whilst the $\mathbf{u}_i$'s form a *finite* set, the right-hand side of (1) will be *infinite* if F is.

Subspaces of V are defined exactly as for $\mathbb{R}^n$:

Definition. *Let V be a vector space over a field F. A subset U of V is a sub*space *iff*

(S1) $\mathbf{u}, \mathbf{v} \in U \ \Rightarrow \ \mathbf{u} + \mathbf{v} \in U$,

(S2) $a \in F,\ \mathbf{u} \in U \ \Rightarrow \ a\mathbf{u} \in U$.

It follows that a subspace is a vector space in its own right: the operations (S1) and (S2) will satisfy all the vector space axioms because V itself does. In practice, subspaces are again defined either by linear combinations or linear equations.

Exercise. Let $V = \mathbb{R}^{3,3}$ be the space of 3×3 matrices. Let $S = \{A \in \mathbb{R}^{3,3} \ : \ A^\top = A\}$ be the subset consisting of *symmetric* matrices. Check that S is a sub*space* of V, and find matrices A_i such that $S = \mathscr{L}\{A_1, \ldots, A_6\}$.

Definition. *A vector space V (for example, a subspace V of some other vector space W) is* finite-dimensional *if it has a* finite *subset* $\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$ *such that* $V = \mathscr{L}\{\mathbf{u}_1, \cdots, \mathbf{u}_k\}$.

Example. Consider $\mathbb{R}^{2,3}$ again. This vector space is finite dimensional because any matrix of size 2×3 is a LC of the matrices

$$\begin{array}{ccc} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, & \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{array} \tag{2}$$

Indeed,

$$\begin{pmatrix} a_5 & a_3 & a_1 \\ a_6 & a_4 & a_2 \end{pmatrix} = a_1\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + a_2\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \cdots + a_6\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{3}$$

Whilst the unordered set (2) is an obvious basis, there is no 'right' or 'wrong' way to order its elements into a list.

The actual dimension of V in the definition above turns out (we shall see) to be the smallest number of vectors needed to 'generate' V, in which case the resulting set is LI. Elements $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in a vector space V are *linearly independent* (LI) if there is no nontrivial linear relation between them. More formally, this means that

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0} \ \ (\text{with } a_i \in F) \quad \Rightarrow \quad a_1 = \cdots = a_k = 0.$$

For example, the six matrices in (2) are LI because (3) can only be null if all the coefficients a_i are zero.

Definition. *A finite set* $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *of elements of a vector space* V *is a* basis *of* V *if*

(B1) it generates V *in the sense that* $V = \mathscr{L}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$*, and*

(B2) it is LI.

Recall that any two bases of a subspace of $\mathbb{R}^n$ have the same number of elements. We shall explain shortly that the same result holds for any finite-dimensional vector space; the dimension of V is then defined to be this number.

Exercise. (i) Guided by (2), show that $\mathbb{R}^{m,n}$ has a basis of consisting of mn matrices.

(ii) Verify that $\{1, x, \dots, x^n\}$ is a basis of $\mathbb{R}_n[x]$, by observing that if $a_0 + a_1x + \cdots + a_nx^n$ equals the *zero polynomial* then it has to vanish for *all* x, and $a_i = 0$ for all i.

L14.2 Linear mappings. Let V, W be two vector spaces. A mapping $f\colon V \to W$ is called *linear* if

(LM1) $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$.

(LM2) $f(a\mathbf{v}) = af(\mathbf{v})$ for all $a \in F$ and $\mathbf{v} \in V$.

These two conditions are equivalent to *either* of the single ones

$$\begin{aligned} f(a\mathbf{u}+b\mathbf{v}) &= af(\mathbf{u}) + bf(\mathbf{v}) \quad \text{for all } a, b \in \mathbb{R},\ \mathbf{u}, \mathbf{v} \in V, \\ f(a\mathbf{u}+\mathbf{v}) &= af(\mathbf{u}) + f(\mathbf{v}) \quad \text{for all } a \in \mathbb{R},\ \mathbf{u}, \mathbf{v} \in V. \end{aligned}$$

Here is an essential consequence:

$$f(\mathbf{0}) = \mathbf{0}, \tag{4}$$

meaning that f maps the null vector of V to the null vector of W (both are denoted here by the symbol $\mathbf{0}$).

Example. Equation (3) effectively defines a linear mapping

$$f\colon \mathbb{R}^{1,6} \to \mathbb{R}^{2,3}, \qquad \text{for which} \quad f(a_1, \dots, a_6) = \begin{pmatrix} a_5 & a_3 & a_1 \\ a_6 & a_4 & a_2 \end{pmatrix}.$$

Here, we have used the notation $f\mathbf{v}$ in place of $f(\mathbf{v})$ to avoid double parentheses. It is easy to check the conditions (LM1) and (LM2); the reason they hold is that each matrix component on the right is a *linear* combination of the coordinates on the left. By contrast, neither of the following mappings is linear:

$$g(a_1, \dots, a_6) = \begin{pmatrix} a_5+1 & a_3 & a_1 \\ a_6 & a_4 & a_2 \end{pmatrix}, \qquad h(a_1, \dots, a_6) = \begin{pmatrix} a_5 & (a_3)^2 & a_1 \\ a_6 & a_4 & a_2 \end{pmatrix}.$$

Let $f\colon \mathscr{A} \to \mathscr{B}$ be an arbitrary mapping between two sets. Recall that the *image* of f,

$$\operatorname{Im} f = \{f(a) : a \in \mathscr{A}\},$$

denotes the subset of $\mathscr{B}$ consisting of those elements 'gotten' from $\mathscr{A}$. Also, given $b \in \mathscr{B}$, its *inverse image*

$$f^{-1}(b) = \{a \in \mathscr{A} : f(a) = b\}$$

is the *subset* of $\mathscr{A}$ consisting of all those elements that map to b. Then f is said to be

(i) *surjective* or *onto* if $\operatorname{Im} f = B$,

(ii) *injective* or *one-to-one* if $f(a_1)=f(a_2) \Rightarrow a_1=a_2$.

Thus f is onto iff $f^{-1}(b)$ is nonempty for all $b \in \mathscr{B}$. If f is both surjective *and* injective then it is called *bijective.* This means that there exists a well-defined inverse mapping

$$f^{-1}\colon \mathscr{B} \to \mathscr{A},$$

so that $f^{-1} \circ f$ and $f \circ f^{-1}$ are identity mappings.

Example. Let n be a positive integer. Then $f(x) = x^n$ defines a bijection $\mathbb{R} \to \mathbb{R}$ iff n is odd; if n is even f is neither injective nor surjective; f is linear only when $n = 1$ (in which case it is the identity mapping). If $n = 2$ then $f^{-1}(64) = \{8, -8\}$; if $n = 3$ then '$f^{-1}(64)$' is ambiguous: it could mean the subset $\{4\}$, or the number 4 obtained by applying the inverse mapping to 64.

Here is an easy but important

Lazy Lemma. *Let $f\colon V \to W$ be a linear mapping. Then f is injective iff*

$$f(\mathbf{v}) = \mathbf{0} \quad \Rightarrow \quad \mathbf{v} = \mathbf{0}. \tag{5}$$

Proof. Equation (4) tells us that (5) is a special case of the injectivity condition. So if f is injective and $f(\mathbf{v}) = \mathbf{0}$, then $f(\mathbf{v}) = f(\mathbf{0})$ and thus $\mathbf{v} = \mathbf{0}$. Conversely, suppose that (5) holds. If $f(\mathbf{v}_1)=f(\mathbf{v}_2)$ then because f is linear,

$$f(\mathbf{v}_2 - \mathbf{v}_1) = f(\mathbf{v}_2) - f(\mathbf{v}_1) = \mathbf{0},$$

and by hypothesis, $\mathbf{v}_2 - \mathbf{v}_1 = \mathbf{0}$. Thus, f is injective. QED

L14.3 Bases and linear mappings. We now use linear mappings to interpret the conditions that define a basis. Suppose that $\mathbf{v}_1, \dots, \mathbf{v}_n$ are any n elements of a vector space V. Define a mapping $f\colon \mathbb{R}^{1,n} \to V$ by

$$(a_1, \dots, a_n) \mapsto a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n. \tag{6}$$

It is easy to check that this mapping is linear. Then (B1) asserts that it is *surjective,* and (B2) implies that it is *injective* (with the help of the Lazy Lemma).

A bijective linear mapping is also called an *isomorphism,* so a basis of V defines an isomorphism f from $\mathbb{R}^n$ to V. Observe from (6) that f maps each element of the canonical basis of $\mathbb{R}^n$ onto an element of the chosen basis of V. If $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is a basis, we may use f to *identify* $\mathbb{R}^n$ with V, and to transfer properties of $\mathbb{R}^n$ to V. This enables one to prove results such as the

Theorem. *Let V be a vector space with a basis of size n. We have*

(i) if m vectors $\mathbf{v}_1, \dots, \mathbf{v}_m$ of V are LI then $m \leqslant n$,

(ii) if $V = \mathscr{L}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ then $n \leqslant p$.

In particular, any *basis of V has n elements, and V is said to have* dimension *n.*

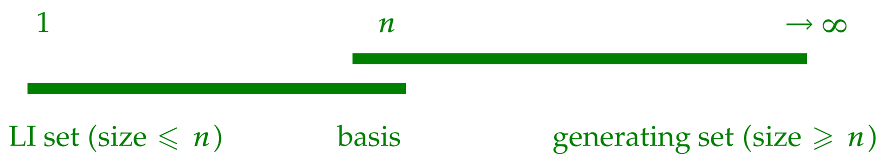
Proof. We already know that this is true for $V = \mathbb{R}^n$. For, we represent the vectors as rows of a matrix A with size $m \times n$ or $p \times n$, and use the theory of the *rank* $r(A)$ of A. Part (i) implies that $r(A) = m$, and so $m \leqslant n$. Part (ii) implies that $r(A) = n$ and so $n \leqslant p$.

To deduce (i) in general, suppose that $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are LI in V. Then $f^{-1}(\mathbf{u}_1), \cdots, f^{-1}(\mathbf{u}_m)$ are LI in $\mathbb{R}^n$ because

$$\mathbf{0} = a_1 f^{-1}(\mathbf{u}_1) + \cdots + a_m f^{-1}(\mathbf{u}_m) = f^{-1}(a_1\mathbf{u}_1 + \cdots + a_m\mathbf{u}_m)$$
$$\Rightarrow \quad \mathbf{0} = a_1\mathbf{u}_1 + \cdots + a_m\mathbf{u}_m \quad \Rightarrow \quad a_1 = \cdots = a_m = 0.$$

Hence $m \leqslant n$. Part (ii) is similar. QED

The statement of the Theorem is represented schematically by



Lazy Corollary. *Suppose that we already know that* $\dim V = n$. *Then in checking whether a set of* n *elements is a basis we only need bother to check* one of *(B1), (B2).*

Proof. A practical way of extending an LI set to a basis is provided by the following result: Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are LI in a vector space V. Then $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}$ are LI iff

$$\mathbf{v}_{k+1} \notin \mathscr{L}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}.$$

We have already seen this in action when $V = \mathbb{R}^n$, and it is valid in general.

Now, if $V = \mathscr{L}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ as in (B1), then the n vectors must be LI. For if not, at least one is redundant and V is generated by $n-1$ elements, impossible. Similarly, if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are LI as in (B2), then they must generate V. For if not, we could add $\mathbf{v}_{n+1}$ to get $n+1$ LI vectors, contradicting the Theorem. QED

L14.4 Further exercises.

1. Let $D\colon \mathbb{R}_3[x] \to \mathbb{R}_3[x]$ denote the mapping given by differentiating: $D(\mathbf{p}(x)) = \mathbf{p}'(x)$. Show that D is linear, and determine the images of the four polynomials $\mathbf{p}_i$ in the Example above. Is D injective? Is it surjective?

2. Consider the polynomials:

$$\mathbf{p}_1(x) = 3 + x^2, \quad \mathbf{p}_2(x) = x - x^2 + 2x^3, \quad \mathbf{p}_3(x) = 2 - x^2, \quad \mathbf{p}_4(x) = x - 2x^2 + 3x^3.$$

Verify that $\{p_1(x), p_2(x), p_3(x), p_4(x)\}$ is a basis of $\mathbb{R}_3[x]$ and express x^2 as a LC of the elements of this basis.

3. Let $A \in \mathbb{R}^{3,3}$ and define $f\colon \mathbb{R}^{3,1} \to \mathbb{R}^{3,1}$ by $f(\mathbf{v}) = A\mathbf{v}$. Prove that f is linear (such examples will be the subject of the next lecture), and use the theory of linear systems to show that f is injective iff $r(A) = 3$.

4. Show that the cubic polynomials

$$\mathbf{p}_1(x) = -\tfrac{1}{6}(x-2)(x-3)(x-4) \qquad \mathbf{p}_2(x) = \tfrac{1}{2}(x-1)(x-3)(x-4)$$
$$\mathbf{p}_3(x) = -\tfrac{1}{2}(x-1)(x-2)(x-4) \qquad \mathbf{p}_4(x) = \tfrac{1}{6}(x-1)(x-2)(x-3)$$

satisfy $\mathbf{p}_i(j) = \delta_{ij}$ (meaning 1 if $i = j$ and 0 otherwise). Deduce that they are LI. Explain carefully why this implies that they form a basis of $\mathbb{R}_3[x]$.