LECTURE 4

Basis sets

The discussion in the previous lecture suggests that minimal spanning sets and maximal linearly independent sets are the same. This is indeed the case.

Theorem 4.1 Let S be a set of vectors in a finite-dimensional vector space $V(\mathbb{F})$. Then S is a minimal spanning set for V if and only if it is a maximal linearly independent set in V.

Proof. The result follows from the characterisations provided by Theorems 3.8 and 3.10. ■

These sets are also known as basis sets. A common definition is included below.

Definition 4.1 — Basis. A basis set for a vector space $V(\mathbb{F})$ is a set S satisfying

- S is a spanning set for $V(\mathbb{F})$;
- S is a linearly independent set.

To establish if a given set is a basis for a vector space, we need to check both the spanning property and the linear independence of the set.

Example 4.1 — Canonical basis for $\mathbb{R}^3$. Let

$$B = \left\{ \mathbf{e}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

It is straightforward to see that B spans $\mathbb{R}^3$ and is a linearly independent set, hence it is a basis for $\mathbb{R}^3$, known as the canonical basis (with the vectors $\mathbf{e}_i$ known as canonical vectors). The generalisation to $\mathbb{R}^n$ is straightforward.

Example 4.2 — Power basis for $\mathcal{P}_2$. Let $B = \{1, x, x^2\}$. Then B is a basis for $\mathcal{P}_2$, known as the power basis. The generalisation to $\mathcal{P}_n$ is straightforward.

By Theorem 3.8, a basis set is a minimal spanning set for $V(\mathbb{F})$. By the discussion in the previous section, we deduce that

- basis sets always exist;
- basis sets are not unique.

The non-uniqueness of a basis set suggests that some bases may be preferred over others: this is indeed the case and we will see later how to identify and construct bases that are convenient in a given context.

The following results describe further properties of basis sets.

Proposition 4.2 Let $V(\mathbb{F})$ be a vector space. Every spanning set for V contains a basis for V.

Proof. Since any basis is a minimal spanning set, the result follows from Corollary 2.10. ■

Proposition 4.3 Let $V(\mathbb{F})$ be a vector space. Every linearly independent set in V is contained in some basis set for V.

Proof. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ be a linearly independent set in $V = \text{span}U$, where $U = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_\ell\}$. We consider the following two complementing cases.
1. Assume that $\mathbf{u}_i \in \text{span}S$ for all $i = 1, 2, \dots, \ell$. Then, $V = \text{span}U \subseteq \text{span}S$. However, since $\mathbf{v}_i \in V$, we have $\text{span}S \subseteq V = \text{span}U$. Therefore $\text{span}S = \text{span}U = V$. This means that S is a linearly independent spanning set for V and therefore a basis. In this case, the result holds trivially since the set S is contained in the basis S.
2. Assume now that $\mathbf{u}_i \notin \text{span}S$ for some i. Then, by Lemma 3.9, $S' := S \cup \{\mathbf{u}_i\}$ is a linearly independent set. If S' is a spanning set for V, then it is a basis and $S \subset S'$, which proves the statement. Otherwise, we replace S with S' and repeat the reasoning in 1 and 2 to obtain a new linearly independent set S' which contains the original set S. If S' is a spanning set, then it is the basis we seek. If not, we continue this procedure: we can do this at most ℓ times; if this indeed happens, we obtain $S' = S \cup U$, in which case S' is a spanning set because U is one. Therefore S' is a basis with $S \subset S'$. ■

The above proof can be used to derive the following corollary.

Corollary 4.4 Let $V(\mathbb{F})$ be a vector space. Let S be a linearly independent set in V and let U be a spanning set for V. Then $|S| \leq |U|$.

These results allow us to derive the following important fact.

Theorem 4.5 Any two basis sets for a finite-dimensional vector space have the same number of elements.

Proof. Let B_1, B_2 be two basis sets for a given vector space. By definition, they are linearly independent spanning sets. This allows us to set $B_1 = U, B_2 = S$ in Corollary 4.4; we deduce that $|B_2| \leq |B_1|$. Reversing the roles of B_1 and B_2, we find $|B_1| \leq |B_2|$, so that $|B_1| = |B_2|$, as stated in the theorem. ■

4.1 Dimension

The result in the previous theorem allows for the following definition.

Definition 4.2 — Dimension. Let $V(\mathbb{F})$ be a non-trivial finite-dimensional vector space. The dimension of V, denoted by $\dim V$, is the number of vectors in a basis for V. If V is trivial, then $\dim V = 0$.

The following results are fairly intuitive.

Proposition 4.6 Let $U \leq V$. Then $\dim U \leq \dim V$, with $\dim U = \dim V$ if and only if $U = V$.

Proof. Let B_U, B_V be basis sets for U and V, respectively. Then B_U is a linearly independent set in V and B_V is a spanning set for V. By Corollary 4.4, $|B_U| \leq |B_V|$ and therefore $\dim U \leq \dim V$.
To show the last statement, let B_U be a basis for U. By Proposition 4.3, we have $B_U \subset B_V$ for some

basis set for V. Since, by hypothesis, dim U = dim V, we have $|B_U| = |B_V|$, so that $B_U = B_V$. Then $U = \text{span}B_U = \text{span}B_V = V$.
On the other hand, if $U = V$, then $U \leq V$ and also $V \leq U$. Hence, by the first statement of the proposition, dim $U \leq$ dim V and also dim $V \leq$ dim U, so that dim U = dim V. ■

Proposition 4.7 Let $V(\mathbb{F})$ be a non-trivial finite-dimensional vector space. Let S be a linearly independent set in V with $|S| = \dim V$. Then S is a basis set for V.

Proof. Since S is linearly independent, it is a basis for span$S$, which is a subspace for V: span$S \leq V$. Since $|S| = \dim V$, by Proposition 4.6, $V = \text{span}S$. Hence S is a spanning set for V, which is also linearly independent, so it is a basis, by definition. ■

We end this section with the following important result.

Theorem 4.8 Let U, V be two subspaces of a finite-dimensional vector space. Then

$$\dim\,(U+V) = \dim\,U + \dim\,V - \dim\,U \cap V.$$

Proof. Let $W := U \cap V$ and define

$$\dim\,U =: k, \quad \dim\,V =: \ell, \quad \dim\,W = m.$$

Let B_W be a basis for W. By Proposition 4.3, since B_W is a linearly independent set in U, it must be contained in some basis B_U of U. Similarly, B_W must be contained in some basis B_V of V. Let us denote these bases as follows:

- $B_W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$,
- $B_U := \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k-m}\}$,
- $B_V := \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{\ell-m}\}$,

where the basis sets B_U, B_V contain elements $\mathbf{u}_j \notin V$ and $\mathbf{v}_s \notin U$, respectively. Then a spanning set for $U + V$ is

$$B := \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k-m}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{\ell-m}\}.$$

This can be readily seen since every element in $U + V$ is a sum of an element of U (hence a linear combination of elements in B_U and an element of V (hence a linear combination of elements in B_V). If B is a linearly independent set, it is a basis set for $U + V$. Therefore,

$$\dim\,(U+V) = |B| = m + (k-m) + (\ell - m) = k + \ell - m = \dim\,U + \dim\,V - \dim\,U \cap V.$$

It remains to show B is indeed linearly independent. Consider the representation of $\mathbf{0}$ as a linear combinations of elements of B:

$$\mathbf{0} = \sum_{i=1}^{m} a_i \mathbf{w}_i + \sum_{j=1}^{k-m} b_j \mathbf{u}_j + \sum_{s=1}^{\ell-m} c_s \mathbf{v}_s.$$

We can re-write this as

$$\sum_{i=1}^{m} a_i \mathbf{w}_i + \sum_{j=1}^{k-m} b_j \mathbf{u}_j = -\sum_{s=1}^{\ell-m} c_s \mathbf{v}_s.$$

The expression on the left represents a vector in U, while that on the right a vector in V. The equality implies that they both represent a vector in both U and V, i.e., in $W = U \cap V$. Since any vector in W is uniquely expressed as a linear combination of $\mathbf{w}_i$, we must have $b_j = 0$ for all j and also $c_s = 0$ for all s. However, this results in another representation of the zero vector involving the elements $\mathbf{w}_i$ of B_W:

$$\sum_{i=1}^{m} a_i \mathbf{w}_i = \mathbf{0}.$$

Since B_W is a linearly independent set, it follows that $a_i =$ for all i. Thus, the intial representation of $\mathbf{0}$ is trivial and the set B is linearly independent. ■

An immediate consequence of this result can be established for the case of direct sums of vector spaces.

Corollary 4.9 Let U, V be two subspaces of a finite-dimensional vector space and define $X = U + V$. Then
$$X = U \oplus V \iff \dim X = \dim U + \dim V.$$

Proof. Let $X = U \oplus V$. By the direct sum criterion 1 we have $U \cap V = \{\mathbf{0}\}$. Hence, since $\dim \operatorname{span}\{\mathbf{0}\} = 0$, we get
$$\dim X = \dim U + \dim V.$$

Conversely, let $X = U + V$ and assume $\dim X = \dim U + \dim V$. By Theorem 4.8, we must have $\dim U \cap V = 0$. Hence, $U \cap V = \{\mathbf{0}\}$ and therefore $X = U \oplus V$. ■

4.2 Coordinates

Recall that given a linearly independent set, we can represent uniquely any vector in its span (see Proposition 3.6). This is an important property which we can employ in the case when the set is a basis for some vector space V. In particular, given a basis $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for a vector space $V(\mathbb{F})$, there exist unique coefficients $x_i \in \mathbb{F}$ such that any $\mathbf{v} \in V(\mathbb{F})$ can be written as

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n.$$

Given their uniqueness, we can provide the following definition.

Definition 4.3 The coefficients $x_1, x_2, \ldots, x_n$ are called the **coordinates** of the vector $\mathbf{v}$ in the basis B.

Note that the coordinates of a vector depend on the choice of basis, which is why sometimes they are also referred to as B-coordinates.

Proposition 4.10 — Coordinate map. Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ denote a basis for V. Let

$$\mathbf{v} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n.$$

Define the **coordinate map** $\varphi : V \mapsto \mathbb{F}^n$ via

$$\varphi(\mathbf{v}) = \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Then φ is a bijection that satisfies the linearity condition

$$\varphi(a\mathbf{x} + b\mathbf{y}) = a\varphi(\mathbf{x}) + b\varphi(\mathbf{y}).$$

R This is an example of a so-called **isomorphism**. We will investigate this topic in Part II when we discuss linear maps. For now, we recall that a bijection, as a one-to-one correspondence, has a unique inverse $\varphi^{-1} : \mathbb{F}^n \mapsto V$: thus, if $\varphi(\mathbf{v}) = \mathbf{x}$, we also have $\varphi^{-1}(\mathbf{x}) = \mathbf{v}$.

Depending on the application of interest, a choice of basis may be preferred over others: often, this choice relates to the resulting coordinates or other evaluations involving basis elements. A suitable basis could yield coordinates that have some physical significance, and/or a sparse set of coordinates, i.e., with only few non-zero values. We will encounter some examples later; for now, we discuss two examples of standard bases.

Example 4.3 — Coordinates in the canonical basis. Let $\mathbb{R}^3$ denote Euclidean space. Recall that the canonical basis was defined to be

$$B = \left\{ \mathbf{e}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{e}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},\ \mathbf{e}_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Then any vector in $\mathbb{R}^3$ can be represented in the following convenient form

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ v_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ v_3 \end{bmatrix} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3.$$

Thus, the coordinates in the canonical basis are simply the entries in the vector $\mathbf{v}$. In Euclidean space, these coordinates have the usual (geometric) significance.

Example 4.4 Let $B = \{1, x, x^2\}$ denote the power basis for $\mathcal{P}_2$. Then any polynomial in $\mathcal{P}_2$ can be represented in the form

$$p(x) = a_0 + a_1 x + a_2 x^2,$$

so that the coordinates in the power basis are just the polynomial coefficients.

We end with an example where a non-standard basis is employed.

Example 4.5 Let

$$B' = \left\{ \mathbf{v}_1 := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},\ \mathbf{v}_2 := \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},\ \mathbf{v}_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Let $\mathbf{v} \in \mathbb{R}^3$ be a generic vector and let us compute its coordinates in the basis B':

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 = \begin{bmatrix} a_1 \\ a_1 + a_2 \\ a_1 + a_2 + a_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \implies \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 - v_1 \\ v_3 - v_2 \end{bmatrix}.$$

Exercise 4.1 Let $B' = \{p_1, p_2, p_3\}$, where $p_i = 1 - (i-1)x^{i-1}$ for $i = 1, 2, 3$. Find the coordinates of the polynomial $p(x) = 1 + 2x + 3x^2$ in the basis B'.