Course 5: 29.10.2020

2.5 Linear independence

Definition 2.5.1 Let V be a vector space over K. We say that the vectors $v_1, \dots, v_n \in V$ are (or the set of vectors $\{v_1, \dots, v_n\}$ is):

(1) *linearly independent* in V if for every $k_1, \dots, k_n \in K$,

$$k_1 v_1 + \dots + k_n v_n = 0 \Longrightarrow k_1 = \dots = k_n = 0\,.$$

(2) *linearly dependent* in V if they are not linearly independent, that is, $\exists k_1, \dots, k_n \in K$ not all zero such that

$$k_1 v_1 + \dots + k_n v_n = 0\,.$$

Remark 2.5.2 (1) A set consisting of a single vector v is linearly dependent $\Longleftrightarrow v = 0$.

(2) As an immediate consequence of the definition, we notice that if V is a vector space over K and $X, Y \subseteq V$ such that $X \subseteq Y$, then:

(i) If Y is linearly independent, then X is linearly independent.

(ii) If X is linearly dependent, then Y is linearly dependent. Thus, every set of vectors containing the zero vector is linearly dependent.

Theorem 2.5.3 *Let V be a vector space over K. Then the vectors $v_1, \dots, v_n \in V$ are linearly dependent if and only if one of the vectors is a linear combination of the others, that is, $\exists j \in \{1, \dots, n\}$ such that*

$$v_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} \alpha_i v_i$$

for some $\alpha_i \in K$, where $i \in \{1, \dots, n\}$ and $i \neq j$.

Proof. $\Longrightarrow$. Assume that $v_1, \dots, v_n \in V$ are linearly dependent. Then $\exists k_1, \dots, k_n \in K$ not all zero, say $k_j \neq 0$, such that $k_1 v_1 + \dots + k_n v_n = 0$. But this implies

$$-k_j v_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} k_i v_i$$

and further,

$$v_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} (-k_j^{-1} k_i) v_i\,.$$

Now choose $\alpha_i = -k_j^{-1} k_i$ for each $i \neq j$ to get the conclusion.

$\Longleftarrow$. Assume that $\exists j \in \{1, \dots, n\}$ such that

$$v_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} \alpha_i v_i$$

for some $\alpha_i \in K$. Then

$$(-1) v_j + \sum_{\substack{i=1 \\ i \neq j}}^{n} \alpha_i v_i = 0\,.$$

Since there exists such a linear combination equal to zero and the scalars are not all zero, the vectors $v_1, \dots, v_n$ are linearly dependent. $\square$

Example 2.5.4 *(a)* Let V_2 be the real vector space of all vectors (in the classical sense) in the plane with a fixed origin O. Recall that the addition is the usual addition of two vectors by the parallelogram rule and the external operation is the usual scalar multiplication of vectors by real scalars. Then:
(i) one vector v is linearly dependent in $V_2 \iff v = 0$;
(ii) two vectors are linearly dependent in $V_2 \iff$ they are collinear;
(iii) three vectors are always linearly dependent in V_2.

Now let V_3 be the real vector space of all vectors (in the classical sense) in the space with a fixed origin O. Then:
(i) one vector v is linearly dependent in $V_3 \iff v = 0$;
(ii) two vectors are linearly dependent in $V_3 \iff$ they are collinear;
(iii) three vectors are linearly dependent in $V_3 \iff$ they are coplanar;
(iv) four vectors are always linearly dependent in V_3.

(b) If K is a field and $n \in \mathbb{N}^*$, then the vectors $e_1 = (1,0,0,\dots,0)$, $e_2 = (0,1,0,\dots,0)$, $\dots$, $e_n = (0,0,0,\dots,1) \in K^n$ are linearly independent in the canonical vector space K^n over K. In order to show that, let $k_1,\dots,k_n \in K$ be such that

$$k_1e_1 + k_2e_2 + \cdots + k_ne_n = 0 \in K^n.$$

Then we have

$$k_1(1,0,0,\dots,0) + k_2(0,1,0,\dots,0) + \cdots + k_n(0,0,0,\dots,1) = (0,\dots,0),$$

and furthermore

$$(k_1,\dots,k_n) = (0,\dots,0).$$

This implies that $k_1 = \cdots = k_n = 0$, and so the vectors $e_1,\dots,e_n$ are linearly independent in K^n.

(c) Let K be a field and $n \in \mathbb{N}$. Then the vectors $1, X, X^2, \dots, X^n$ are linearly independent in the vector space $K_n[X] = \{f \in K[X] \mid degree(f) \leq n\}$ over K.

Let us now give a very useful practical result on linear dependence.

Theorem 2.5.5 *Let $n \in \mathbb{N}$, $n \geq 2$.*
(i) Two vectors in the canonical vector space K^n are linearly dependent $\iff$ their components are respectively proportional.
(ii) n vectors in the canonical vector space K^n are linearly dependent $\iff$ the determinant consisting of their components is zero.

Proof. *(i)* Let $v = (x_1,\dots,x_n)$, $v' = (x'_1,\dots,x'_n) \in K^n$. By Theorem 2.5.3, the vectors v and v' are linearly dependent if and only if one of them is a linear combination of the other, say $v' = kv$ for some $k \in K$. That is, $x'_i = kx_i$ for each $i \in \{1,\dots,n\}$.

(ii) Let $v_1 = (x_{11}, x_{21}, \dots, x_{n1})$, $\dots$, $v_n = (x_{1n}, x_{2n}, \dots, x_{nn}) \in K^n$. The vectors $v_1,\dots,v_n$ are linearly dependent if and only if $\exists k_1,\dots,k_n \in K$ not all zero such that

$$k_1v_1 + \cdots + k_nv_n = 0\,.$$

But this is equivalent to

$$k_1(x_{11}, x_{21}, \dots, x_{n1}) + \cdots + k_n(x_{1n}, x_{2n}, \dots, x_{nn}) = (0,\dots,0)\,,$$

and further to

$$\begin{cases} k_1x_{11} + k_2x_{12} + \cdots + k_nx_{1n} = 0 \\ k_1x_{21} + k_2x_{22} + \cdots + k_nx_{2n} = 0 \\ \dots\dots\dots\dots\dots\dots \\ k_1x_{n1} + k_2x_{n2} + \cdots + k_nx_{nn} = 0\,. \end{cases}$$

We are interested in the existence of a non-zero solution for this homogeneous linear system. We will see later on that such a solution does exist if and only if the determinant of the system is zero. □

2.6 Basis

We are going to define a key notion related to a vector space, namely that of a *basis*, which will perfectly determine a vector space. For the sake of simplicity and because of our limited needs, til the end of the chapter, *by a vector space we will understand a finitely generated vector space.*

Definition 2.6.1 Let V be a vector space over K. A list of vectors $B = (v_1, \dots, v_n) \in V^n$ is called a *basis* of V if:
(1) B is linearly independent in V;
(2) B is a system of generators for V, that is, $\langle B \rangle = V$.

Theorem 2.6.2 *Every vector space has a basis.*

Proof. Let V be a vector space over K. If $V = \{0\}$, then it has the basis $\emptyset$.

Now let $V = \langle B \rangle \neq \{0\}$, where $B = (v_1, \dots, v_n)$. If B is linearly independent, then B is a basis and we are done. Suppose that the list B is linearly dependent. Then by Theorem 2.5.3, $\exists j_1 \in \{1, \dots, n\}$ such that

$$v_{j_1} = \sum_{\substack{i=1 \\ i \neq j_1}}^{n} k_i v_i$$

for some $k_i \in K$. It follows that $V = \langle B \setminus \{v_{j_1}\} \rangle$, because every vector of V can be written as a linear combination of the vectors of $B \setminus \{v_{j_1}\}$. If $B \setminus \{v_{j_1}\}$ is linearly independent, it is a basis and we are done. Otherwise, $\exists j_2 \in \{1, \dots, n\} \setminus \{j_1\}$ such that

$$v_{j_2} = \sum_{\substack{i=1 \\ i \neq j_1, j_2}}^{n} k'_i v_i$$

for some $k'_i \in K$. It follows that $V = \langle B \setminus \{v_{j_1}, v_{j_2}\} \rangle$, because every vector of V can be written as a linear combination of the vectors of $B \setminus \{v_{j_1}, v_{j_2}\}$. If $B \setminus \{v_{j_1}, v_{j_2}\}$ is linearly independent, then it is a basis and we are done. Otherwise, we continue the procedure. If all the previous intermediate subsets are linearly dependent, we get to the step $V = \langle B \setminus \{v_{j_1}, \dots, v_{j_{n-1}}\} \rangle = \langle v_{j_n} \rangle$. If v_{j_n} were linearly dependent, then $v_{j_n} = 0$, hence $V = \langle v_{j_n} \rangle = \{0\}$, contradiction. Hence v_{j_n} is linearly independent and thus forms a single element basis of V. □

Remark 2.6.3 We are going to see that a vector space may have more than one basis.

Let us give now a characterization theorem for a basis of a vector space.

Theorem 2.6.4 *Let V be a vector space over K. A list $B = (v_1, \dots, v_n)$ of vectors in V is a basis of V if and only if every vector $v \in V$ can be uniquely written as a linear combination of the vectors $v_1, \dots, v_n$, that is,*

$$v = k_1 v_1 + \dots + k_n v_n$$

for some unique $k_1, \dots, k_n \in K$.

Proof. $\Longrightarrow$. Assume that B is a basis of V. Hence B is linearly independent and $\langle B \rangle = V$. The second condition assures us that every vector $v \in V$ can be written as a linear combination of the vectors of B. Suppose now that $v = k_1 v_1 + \dots + k_n v_n$ and $v = k'_1 v_1 + \dots + k'_n v_n$ for some $k_1, \dots, k_n, k'_1, \dots, k'_n \in K$. It follows that

$$(k_1 - k'_1)v_1 + \dots + (k_n - k'_n)v_n = 0\,.$$

By the linear independence of B, we must have $k_i = k'_i$ for each $i \in \{1, \dots, n\}$. Thus, we have proved the uniqueness of writing.

$\Longleftarrow$. Assume that every vector $v \in V$ can be uniquely written as a linear combination of the vectors of B. Then clearly, $V = \langle B \rangle$. For $k_1, \dots, k_n \in K$, we have by the uniqueness of writing

$$k_1 v_1 + \dots + k_n v_n = 0 \Longrightarrow k_1 v_1 + \dots + k_n v_n = 0 \cdot v_1 + \dots + 0 \cdot v_n \Longrightarrow$$

$$\Longrightarrow k_1 = \dots = k_n = 0\,,$$

hence B is linearly independent. Consequently, B is a basis of V. □

Definition 2.6.5 Let V be a vector space over K, $B = (v_1, \dots, v_n)$ a basis of V and $v \in V$. Then the scalars $k_1, \dots, k_n \in K$ intervening in the unique writing of v as a linear combination

$$v = k_1 v_1 + \cdots + k_n v_n$$

of the vectors of B are called the *coordinates of v in the basis B*.

Example 2.6.6 (a) If K is a field and $n \in \mathbb{N}^*$, then the list $E = (e_1, \dots, e_n)$ of vectors of K^n, where

$$\begin{cases} e_1 = (1, 0, 0, \dots, 0) \\ e_2 = (0, 1, 0, \dots, 0) \\ \dots\dots\dots \\ e_n = (0, 0, 0, \dots, 1) \end{cases}$$

is a basis of the canonical vector space K^n over K, called the *canonical basis*. Indeed, each vector $v = (x_1, \dots, x_n) \in K^n$ has a unique writing $v = x_1 e_1 + \cdots + x_n e_n$ as a linear combination of the vectors of E, hence E is a basis of V by Theorem 2.6.4.

Notice that the coordinates of a vector in the canonical basis are just the components of that vector, fact that is not true in general.

(b) Consider the canonical real vector space $\mathbb{R}^2$. We already know a basis of $\mathbb{R}^2$, namely the canonical basis $((1, 0), (0, 1))$. But it is easy to show that the list $((1, 1), (0, 1))$ is also a basis of $\mathbb{R}^2$. Therefore, a vector space may have more than one basis.

(c) Let V_3 be the real vector space of all vectors (in the classical sense) in the space with a fixed origin O. Then a basis of V_3 consists of the three pairwise orthogonal *unit vectors* $\overrightarrow{i}$, $\overrightarrow{j}$, $\overrightarrow{k}$.

(d) Let K be a field and $n \in \mathbb{N}$. Then the list $B = (1, X, X^2, \dots, X^n)$ is a basis of the vector space $K_n[X] = \{f \in K[X] \mid degree(f) \leq n\}$ over K, because every vector (polynomial) $f \in K_n[X]$ can be uniquely written as a linear combination $a_0 \cdot 1 + a_1 \cdot X + \cdots + a_n \cdot X^n$ $(a_0, \dots, a_n \in K)$ of the vectors of B (see Theorem 2.6.4).

In this case, the coordinates of a vector $f \in K_n[X]$ in the basis B are just its coefficients as a polynomial.

(e) Let K be a field. The list

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

is a basis of the vector space $M_2(K)$ over K.

More generally, let $m, n \in \mathbb{N}$, $m, n \geq 2$ and consider the matrices $E_{ij} = (a_{kl})$, where

$$a_{kl} = \begin{cases} 1 & \text{if } k = i \text{ and } l = j \\ 0 & \text{otherwise} \end{cases}.$$

Then the list consisting of all matrices E_{ij} is a basis of the vector space $M_{mn}(K)$ over K.

In this case, the coordinates of a vector $A \in M_{mn}(K)$ in the above basis are just the entries of that matrix.

Theorem 2.6.7 *Let $f : V \to V'$ be a K-linear map and let $B = (v_1, \dots, v_n)$ be a basis of V. Then f is determined by its values on the vectors of the basis B.*

Proof. Let $v \in V$. Since B is a basis of V, $\exists! k_1, \dots, k_n \in K$ such that $v = k_1 v_1 + \cdots + k_n v_n$. Then

$$f(v) = f(k_1 v_1 + \cdots + k_n v_n) = k_1 f(v_1) + \cdots + k_n f(v_n),$$

that is, f is determined by $f(v_1)$, ..., $f(v_n)$. □

Corollary 2.6.8 *Let $f, g : V \to V'$ be K-linear maps and let $B = (v_1, \dots, v_n)$ be a basis of V. If $f(v_i) = g(v_i)$, $\forall i \in \{1, \dots, n\}$, then $f = g$.*

Proof. Let $v \in V$. Then $v = k_1 v_1 + \cdots + k_n v_n$ for some $k_1, \dots, k_n \in K$, hence

$$f(v) = f(k_1 v_1 + \cdots + k_n v_n) = k_1 f(v_1) + \cdots + k_n f(v_n) = k_1 g(v_1) + \cdots + k_n g(v_n) = g(v).$$

Therefore, $f = g$. □

Extra: Lossy compression

Definition 2.6.9 Let $k, n \in \mathbb{N}^*$ be such that $k < n$, and let u be a vector of the canonical vector space K^n over K. Then the *closest k-sparse* vector associated to u is defined as the vector obtained from u by replacing all but its k largest magnitude components by zero.

Example 2.6.10 Consider an image consisting of a single row of four pixels with intensities 200, 50, 200 and 75 respectively. We know that such an image can be viewed as a vector $u = (200, 50, 200, 75)$ in the real canonical vector space $\mathbb{R}^4$. The closest 2-sparse vector associated to u is the vector $\tilde{u} = (200, 0, 200, 0)$.

Suppose that we need to store a grayscale image of (say) $n = 2000 \times 1000$ pixels more compactly. We can view it as a vector v in the real canonical vector space $\mathbb{R}^n$. If we just store its associated closest k-sparse vector, then the compressed image may be far from the original.

One may use the following *lossy compression algorithm*:

Step 1. *Consider a* suitable *basis $B = (v_1, \dots, v_n)$ of the real canonical vector space $\mathbb{R}^n$.*

Step 2. *Determine the n-tuple u (which is desired to have as many zeros as possible) of the coordinates of v in the basis B.*

Step 3. *Replace u by the closest k-sparse n-tuple $\tilde{u}$ for a* suitable *k, and store $\tilde{u}$.*

Step 4. *In order to recover an image from $\tilde{u}$, compute the corresponding linear combination of the vectors of B with scalars the components of $\tilde{u}$.*

Consider the following image:

First, use the closest sparse vector which supresses all but 10% of the components of v, and secondly, use the lossy compression algorithm which supresses all but 10% of the components of u in order to get the following images respectively:

Reference: P.N. Klein, Coding the Matrix. Linear Algebra through Applications to Computer Science*, Newtonian Press, 2013.*