Linear Combinations and Span.

Definition. Let V be a vector space over a field F, and let $u_1, \dots, u_n$ be vectors in V.
A vector $v \in V$ is a *linear combination* of the vectors $u_1, \dots, u_n$ if there exist scalars $a_1, \dots, a_n$ in F such that $v = a_1 u_1 + a_2 u_2 + \dots + a_n u_n$.
We refer to $a_1, \dots, a_n$ as the *coefficients* of the linear combination.

- In any vector space V, we always have $0v = 0$ for each $v \in V$. Thus the zero vector is a linear combination of any non-empty set of vectors in V.
- Given $v \in V$ and $u_1, \dots, u_n \in V$, how can one determine whether v is a linear combination of the vectors $u_1, \dots, u_n$?

That is, we need to understand if it is possible to find scalars $a_1, \dots, a_n \in F$ such that $a_1 u_1 + \dots + a_n u_n = v$.
This question often reduces to solving a *system of linear equations*.

Example. Let $V = \mathbb{R}^2$, let $v = (1,5)$ and $u_1 = (2,0)$, $u_2 = (3,-1)$.
We must determine whether there are scalars $a_1, a_2 \in \mathbb{R}$ such that $v = a_1 u_1 + a_2 u_2$, or:
$(1,5) = a_1(2,0) + a_2(3,-1) = (2a_1, 0) + (3a_2, -a_2) = (2a_1 + 3a_2, -a_2)$.
So v is a linear combination of u_1, u_2 iff there are real numbers $a_1, a_2 \in \mathbb{R}$ such that the system of linear equations

(1) $1 = 2a_1 + 3a_2$
(2) $5 = -a_2$

is satisfied.
Note that then necessarily $a_2 = -5$ by (2), hence $1 = 2a_1 + 3(-5)$, so $2a_1 = 16$, so $a_1 = 8$.
Then $(a_1, a_2) = (8, -5)$ is a solution, showing that indeed v is a linear combination of u_1, u_2.
(For a more involved example see Textbook, Section 1.4 and Problem Set 2).

- Theorem 1.3 allows us to determine when a subset S of a vector space V is already a subspace.

Now we discuss how, starting with an arbitrary subset $S \subseteq V$ (which may not be a subspace of V itself), to find a subspace of V "generated" by it.

Definition. Let S be any subset of a vector space V.
The *span of S*, denoted $\mathrm{Span}(S)$, is the set consisting of all linear combinations of the vectors in S. That is,
$\mathrm{Span}(S) = \{a_1 u_1 + \dots + a_n u_n : n \in \mathbb{N}, a_i \in F, u_i \in S\} \subseteq V$.

As the empty set $\emptyset$ is a subset of any vector space V, $\mathrm{Span}(\emptyset)$ also needs to be defined.
For convenience, we define $\mathrm{Span}(\emptyset) = \{0\}$.
Note that $S \subseteq \mathrm{Span}(S)$, since for every $u \in X$, $u = 1 \cdot u \in \mathrm{Span}(S)$.

Example. Consider the vectors $u_1 = (1,0,0)$, $u_2 = (0,1,0)$ in $\mathbb{R}^3$. Let $S = \{u_1, u_2\} \subseteq V$.
Then vectors in $\mathrm{Span}(S)$ are precisely the vectors of the form $a_1 u_1 + a_2 u_2$ where a_1, a_2 vary over $\mathbb{R}$.
That is, $\mathrm{Span}(S)$ consists of all the vectors of the form $a_1(1,0,0) + a_2(0,1,0) = (a_1, a_2, 0)$ for some $a_1, a_2 \in \mathbb{R}$.
Thus $\mathrm{Span}(S) = \{(a_1, a_2, 0) : a_1, a_2 \in \mathbb{R}\}$ — we already know that this is a subspace of V.
This is not a coincidence!

Theorem 1.5. Let V be a vector space over F.
1) The span of any subset S of V is a subspace of V.
2) Any subspace of V that contains S must also contain $\mathrm{Span}(S)$.
(so $\mathrm{Span}(S)$ is the smallest subspace of V that contains S)

Proof. Both (1) and (2) are obvious if $S = \emptyset$, because $\mathrm{Span}(\emptyset) = \{0\}$ — we know that it is a subspace of V, and any subspace of V must contain 0.
If $S \neq \emptyset$, then S contains a vector z. As $0z = 0$, $0 \in \mathrm{Span}(S)$.
Let $x, y \in \mathrm{Span}(S)$. Then we can write
$x = a_1 u_1 + a_2 u_2 + \dots + a_m u_m$ and $y = b_1 v_1 + \dots + b_n v_n$ for some $a_1, \dots, a_n, b_1, \dots, b_n \in F$ and $u_1, \dots, u_m, v_1, \dots, v_n \in S$.

Then both
$x+y = a_1u_1 + \dots + a_m u_m + b_1 v_1 + \dots + b_n v_n$ and $c\cdot x = (ca_1)u_1 + (ca_2)u_2 + \dots + (ca_m)u_m$
are also linear combinations of vectors of S, and so belong to Span (S).
Thus (a),(b),(c) in Theorem 1.3 hold and it follows that Span(C) is a subspace of V, showing (1).
Now let $W \subseteq V$ be any subspace of V that contains S.
If $w \in$ Span(S) then we can write
$w = c_1 w_1 + \dots + c_k w_k$
for some vectors $w_1, \dots, w_k \in S$ and some scalars $c_1, \dots, c_k$ in F.
Since $S \subseteq W$, we have $w_1, \dots, w_k \in W$.
But as W is a vector space, this implies that $w = c_1w_1 + \dots + c_k w_k$ is also in W. (W is closed under addition and multiplication)
Because w, an arbitrary vector in Span(S), belongs to W, it follows that Span$(S) \subseteq W$.
This proves (2).

Definition. A subset S of a vector space V **generates** (or **spans**) V if Span$(S) = V$.
(In this case, we also say that the vectors of S generate, or span, V.)

- Finding a small $\underbrace{\text{generating set}}_{}$ (and explicit) for a vector space is an efficient way of describing V and simplifies working with it.

Example For any vector space V, Span$(V) = V$ (so V is generated by itself).

Example. The vectors $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ generate the vector space $\mathbb{R}^3$.
Indeed, any vector $(a,b,c) \in \mathbb{R}^3$ can be written as $a\cdot(1,0,0) + b(0,1,0) + c(0,0,1)$, and so
Span $(\{(1,0,0), (0,1,0), (0,0,1)\}) = \mathbb{R}^3$.

Example. Let $V = M_{2\times2}(\mathbb{R})$ be the vector space of all 2×2 matrices with entries from $\mathbb{R}$.
Then $M_1 = \begin{pmatrix}1&0\\0&0\end{pmatrix}$, $M_2 = \begin{pmatrix}0&1\\0&0\end{pmatrix}$, $M_3 = \begin{pmatrix}0&0\\1&0\end{pmatrix}$ and $M_4 = \begin{pmatrix}0&0\\0&1\end{pmatrix}$ generate $M_{2\times2}(\mathbb{R})$.
Indeed, for any $a,b,c,d \in \mathbb{R}$ we can write

$$\begin{pmatrix}a&b\\c&d\end{pmatrix} = a\begin{pmatrix}1&0\\0&0\end{pmatrix} + b\begin{pmatrix}0&1\\0&0\end{pmatrix} + c\begin{pmatrix}0&0\\1&0\end{pmatrix} + d\begin{pmatrix}0&0\\0&1\end{pmatrix}.$$

Thus Span $(\{M_1, \dots, M_4\}) = M_{2\times2}(\mathbb{R})$.

Example. Let $P(F)$ be the vector space of all polynomials over F.
Then the set $\{1, x, x^2, x^3, \dots\}$ generates $P(F)$.
Indeed, Span $(\{1, x, x^2, \dots\}) = \{a_0 + a_1x + \dots + a_nx^n : n \in \mathbb{N}, a_i \in F\}$ — all polynomials over F appear.
Similarly, $P_n(F)$ is generated by $\{1, x, \dots, x^n\}$.

Linear Independence

Usually there are many subsets that generate the same space.

Example.
We saw that the vectors $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ generate the vector space $\mathbb{R}^3$.
The set $\{(1,0,0), (0,1,0), (0,0,1), (2,3,-1)\}$ also generates $\mathbb{R}^3$, but in fact the vector $(2,3,-1)$ is redundant.

- It is natural to look for the smallest possible subset of V that generates it.
First, we explore the circumstances under which a vector can be removed from a generating set to obtain a smaller generating set.

- If $u_1, \dots, u_n$ are any vectors in a vector space V over F, then the zero vector is always a linear combination of $u_1, \dots, u_n$:
$0 = 0\cdot u_1 + 0\cdot u_2 + \dots + 0\cdot u_n$.

(a linear combination is trivial if all of the coefficients $a_1, \dots, a_n$ are equal to zero).

- Sometimes it is also possible to write

$0 = a_1 u_1 + \dots + a_n u_n$

so that not all of $a_1, \dots, a_n \in F$ are 0. (so 0 is a non-trivial linear combination of $u_1, \dots, u_n$).

Example. In $\mathbb{R}^2$, $0 = 2\cdot(1,2) + 5(2,1) + 3(-4,-3)$ is a non-trivial representation of 0.

Definition. A subset S of a vector space V is called linearly dependent if there exist a finite number of distinct vectors $u_1 \neq u_2 \neq \dots \neq u_n$ in S and scalars $a_1, \dots, a_n \in F$, not all zero, such that $a_1 u_1 + a_2 u_2 + \dots + a_n u_n = 0$.
If S is not linearly dependent, then S is called linearly independent.
(We also say that the vectors $v_1, \dots, v_n$ are linearly dependent / independent if the set $\{v_1, \dots, v_n\}$ is linearly dependent / independent.)

Example. Let $V = \mathbb{R}^2$.

- The set $S_1 = \{(0,1), (1,0)\}$ is linearly independent.
 Indeed, if $(0,0) = a_1(0,1) + a_2(1,0)$, then $\begin{cases} 0 = a_1\cdot 0 + a_2 \cdot 1 \\ 0 = a_1 \cdot 1 + a_2 \cdot 0 \end{cases}$ must hold, and so $a_1 = 0$, $a_2 = 0$. This means that any representation of the zero vector as a linear combination of vectors from S_1 is trivial.
- However, the set $S_2 = \{(0,1), (1,0), (17,18)\}$ is linearly dependent.
 Indeed, $18(0,1) + 17(1,0) + (-1)(17,18) = 0$, and this is a non-trivial representation of 0.

Example. V be an arbitrary vector space over F.

1) Any set $S \subseteq V$ containing 0 is linearly dependent. (indeed, as $0 \in S$, then $0 = 1 \cdot 0$ is a non-trivial representation of 0 as a lin. comb. of vectors in V).
2) The empty set $\emptyset \subseteq V$ is linearly independent (we cannot form any linear combination at all using its elements)
3) If $S = \{u\} \subseteq V$ consists of a single non-zero vector u, then S is linearly independent.
 Indeed, if $\{u\}$ is linearly dependent, then $au = 0$ for some non-zero scalar $a \in F$. Thus
 $u = \underbrace{(a^{-1} \cdot a)}_{=1} u = a^{-1}(au) = a^{-1} \cdot 0 = 0$ (by Theorem 1.2)

Example. Similarly, in $V = M_{2\times2}(\mathbb{R})$, the set $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ is linearly independent.
Indeed, assume that
$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = a_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + a_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + a_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ — a representation of the zero vector $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in M_{2\times2}(\mathbb{R})$.
This means that the system of linear equations

$$\begin{cases} 0 = a_1 \cdot 1 + \cancel{a_2 \cdot 0} + \cancel{a_3 \cdot 0} + \cancel{a_4 \cdot 0} \\ 0 = \cancel{a_1 \cdot 0} + a_2 \cdot 1 + \cancel{a_3 \cdot 0} + \cancel{a_4 \cdot 0} \\ 0 = \cancel{a_1 \cdot 0} + \cancel{a_2 \cdot 0} + a_3 \cdot 1 + \cancel{a_4 \cdot 0} \\ 0 = \cancel{a_1 \cdot 0} + \cancel{a_2 \cdot 0} + \cancel{a_3 \cdot 0} + a_4 \cdot 1 \end{cases}$$

is satisfied. But this is only possible when $a_1 = a_2 = a_3 = a_4 = 0$. Thus, there are no non-trivial representation of 0 using elements from S.

Example. In $V = P_n(F)$, the set $S = \{1, x, \dots, x^n\}$ is linearly independent.
Indeed, assume that
$0 = a_0 \cdot 1 + a_1 \cdot x + \dots + a_n x^n$ is a representation of the zero vector in $P_n(F)$ (which is the zero polynomial).
This is only possible when all of a_i, $i = 0, \dots, n$ are 0, which implies that the representation is trivial.

Theorem 1.6. Let V be a vector space over F, and let $S_1 \subseteq S_2 \subseteq V$ be two subsets.
1) S_1 is linearly dependent $\Rightarrow$ S_2 is linearly dependent.
2) S_2 is linearly independent $\Rightarrow$ S_1 is linearly independent.

Proof. (2) follows from (1).

To see (i), notice that if $a_1u_1 + \dots + a_nu_n$ is a linear combination of elements of S_1, then it is also a linear combination of elements of S_2.

Now we connect the notions of span and linear independence.

Theorem 1.7. Let S be a linearly independent subset of a vector space V, and let v be a vector in V that is not in S.
Then the set $S \cup \{v\}$ is linearly dependent if and only if $v \in \mathrm{Span}(S)$.

Proof. ($\Rightarrow$)
If the set $S \cup \{v\}$ is linearly dependent then we can write
$$0 = a_1u_1 + \dots + a_nu_n$$
for some $u_1, \dots, u_n \in S \cup \{v\}$ and some non-zero scalars $a_1, \dots, a_n \in F$.
Because S is linearly independent, it is not possible that $u_i \in S$ for all $i = 1, \dots, n$.
Thus one of the u_i, let's say u_1, equals v. Then $a_1v + a_2u_2 + \dots + a_nu_n = 0$. As $a_1 \neq 0$, we get:
$$v = \frac{1}{a_1}\left(-a_2u_2 - \dots - a_nu_n\right) = \left(-\frac{a_2}{a_1}\right)u_2 + \left(-\frac{a_3}{a_1}\right)u_3 + \dots + \left(-\frac{a_n}{a_1}\right)u_n.$$

As $\left(-\frac{a_i}{a_1}\right)$ are scalars in F for all $i = 2, \dots, n$, it follows that v is a linear combination of the vectors $u_2, \dots, u_n \in S$. Thus $v \in \mathrm{Span}(S)$.

($\Leftarrow$)
Conversely, let $v \in \mathrm{Span}(S)$. Then we can write
$$v = b_1v_1 + \dots + b_mv_m$$
for some vectors $v_1, \dots, v_m \in S$ and some scalars $b_1, \dots, b_m \in F$. Hence
$$0 = b_1v_1 + \dots + b_mv_m + (-1)v.$$
Since v is not in S, in particular $v \neq v_i$ for all $i = 1, \dots, m$, and all of $v_1, \dots, v_m, v$ are distinct.
As the coefficient of v in this linear combination is non-zero $(=-1)$, the set $\{v_1, \dots, v_m, v\}$ is linearly dependent.
As $\{v_1, \dots, v_m, v\} \subseteq S \cup \{v\}$, the set $S \cup \{v\}$ is linearly dependent by Theorem 1.6.

Corollary If S generates a subspace W and no proper subset of S generates W, then S is linearly indep.

Bases and dimension

Definition. A **basis** β for a vector space V is a linearly independent subset of V that generates V.

Example 1. Recall that $\mathrm{Span}(\emptyset) = \{0\}$ and $\emptyset$ is linearly independent. Thus the empty set $\emptyset$ is a basis for the zero vector space.

Example 2. In $V = F^n$, let $e_1 = (1, 0, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, $\dots$, $e_n = (0, 0, \dots, 0, 1)$.
Then $\{e_1, e_2, \dots, e_n\}$ is a basis for F^n and is called the standard basis for F^n.
(In the previous examples we have already checked this for $\mathbb{R}^2$).

Example 3. In $V = M_{m\times n}(F)$, let E^{ij} denote the matrix whose only nonzero entry is a 1 in the i^{th} row and j^{th} column.
Then the set $\{E^{ij} : 1 \le i \le m, 1 \le j \le n\}$ is a basis for $M_{m\times n}(F)$.
(Again, in the previous examples we have already checked this for $M_{2\times 2}(\mathbb{R})$.)

$$E^{ij} = \begin{array}{c} \\ \\ i \\ \\ \end{array} \begin{array}{c} \quad\quad j \\ \left(\begin{array}{c|c} 0 & 0 \\ \hline & \\[-8pt] \!\!\!\!\!\!\mathbf{1} & \\ 0 & 0 \end{array}\right) \end{array}$$

Example 4. In $V = P_n(F)$, the set $\{1, x, x^2, \dots, x^n\}$ is a basis. We call this basis the standard basis for $P_n(F)$.

Example 5. In $V = P(F)$, the set $\{1, x, x^2, x^3, \dots\}$ is a basis.
This shows in particular that a basis need not be finite.
(Later we will see that in fact no basis for $P(F)$ can be finite.)

The next theorem establishes the most significant property of a basis:
- Every vector in V can be expressed in one and only one way as a linear combination of the vectors in the basis.

It is this property that makes bases the building blocks of vector spaces.

Theorem 1.8. Let V be a vector space and $\beta = \{u_1, \dots, u_n\}$ be a subset of V. Then the following two statements are equivalent.

1) β is a basis for V
2) Every vector $v \in V$ can be uniquely expressed as a linear combination of vectors in β, that is, can be expressed in the form $v = a_1u_1 + \dots + a_nu_n$ for unique scalars $a_1, \dots, a_n$.

Proof. (1) implies (2).

Let β be a basis for V. If $v \in V$ is any vector in V, then $v \in \text{Span}(\beta)$ because $\text{Span}(\beta) = V$ (by assumption). Thus v is a linear combination of the vectors in β.
Suppose that

$v = a_1u_1 + a_2u_2 + \dots + a_nu_n$ and $v = b_1u_1 + b_2u_2 + \dots + b_nu_n$

are any two such representations of v.
Subtracting the 2^{nd} equation from the 1^{st} one and rearranging gives:

$0 = (a_1 - b_1)u_1 + (a_2 - b_2)u_2 + \dots + (a_n - b_n)u_n$.

Since β is linearly independent, it follows that

$a_1 - b_1 = 0, \dots, a_n - b_n = 0$ (otherwise we would get a non-trivial representation of 0 with vectors in β).

Hence $a_1 = b_1, \dots, a_n = b_n$ — which means that there is a unique way to express v as a lin. comb. of the vectors in β.

(2) implies (1).

Suppose every $v \in V$ can be uniquely expressed as a lin. comb. of $u_1, \dots, u_n$.
Then $\text{Span}(\beta) = V$ (in particular), and it remains to check that β is linearly independent.
Assume $0 = a_1u_1 + \dots + a_nu_n$ for some $a_1, \dots, a_n$ in F.
But also $0 = 0 \cdot u_1 + \dots + 0 \cdot u_n$.
These are two ways to express $0 \in V$, so by the uniqueness assumption they must coincide.
That is, we must have $a_1 = 0$, $a_2 = 0$, $\dots$, $a_n = 0$ — which means that β is lin. indep.