Optimization and Computational Linear Algebra for Data Science
Lecture 1: Vector spaces

Léo Miolane · leo.miolane@gmail.com

July 10, 2019

Warning: *This material is not meant to be lecture notes. It only gathers the main concepts and results from the lecture, without any additional explanation, motivation, examples, figures...*

1 General definitions

We present below the abstract mathematical definition of a vector space. **Please do not try to memorize it!** Simply remember that a vector space is a set whose elements are called *vectors*, that one can add vectors together and multiply them by real numbers called *scalars*.

Definition 1.1 (*Vector space*)

A vector space (over $\mathbb{R}$) consists of of a set V (whose elements are called vectors) and two operations $+$ and $\cdot$ that verify:

1. *The sum of two vectors is a vector: for all $\vec{x}, \vec{y} \in V$ we have $\vec{x} + \vec{y} \in V$.*
2. *The vector sum is commutative and associative. For all $\vec{x}, \vec{y}, \vec{z} \in V$ we have*
$$\vec{x} + \vec{y} = \vec{y} + \vec{x} \qquad \text{and} \qquad \vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}.$$
3. *There exists a zero vector $\vec{0} \in V$ that verifies $\vec{x} + \vec{0} = \vec{x}$ for all $\vec{x} \in V$.*
4. *For all $\vec{x} \in V$, there exists $\vec{y} \in V$ such that $\vec{x} + \vec{y} = \vec{0}$. Such $\vec{y}$ is called the additive inverse of $\vec{x}$ and is written $-\vec{x}$.*
5. *Scalar multiplication: for all $\vec{x} \in V$ and all $\alpha \in \mathbb{R}$, $\alpha \cdot \vec{x} \in V$.*
6. *Identity element for scalar multiplication: $1 \cdot \vec{x} = \vec{x}$ for all $\vec{x} \in V$.*
7. *Compatibility between scalar multiplication and the usual multiplication: for all $\alpha, \beta \in \mathbb{R}$ and all $\vec{x} \in V$, we have*
$$\alpha \cdot (\beta \cdot \vec{x}) = (\alpha\beta) \cdot \vec{x}.$$
8. *Distributivity: for all $\alpha, \beta \in \mathbb{R}$ and all $\vec{x}, \vec{y} \in V$,*
$$(\alpha + \beta) \cdot \vec{x} = \alpha \cdot \vec{x} + \beta \cdot \vec{y} \qquad \text{and} \qquad \alpha \cdot (\vec{x} + \vec{y}) = \alpha \cdot \vec{x} + \alpha \cdot \vec{y}.$$

From now we will ignore $\cdot$ and simply write $\alpha\vec{x}$ instead of $\alpha \cdot \vec{x}$.

Example 1.1.

- The set $V = \mathbb{R}^n$ endowed with the usual vector addition $+$
$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$
and the usual scalar multiplication $\cdot$
$$\alpha \cdot (x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n)$$
is a vector space.

- The set $V = \mathcal{F}(\mathbb{R},\mathbb{R}) \stackrel{\text{def}}{=} \{f \mid f : \mathbb{R} \to \mathbb{R}\}$ of all functions from $\mathbb{R}$ to itself endowed with the addition $+$ and the scalar multiplication $\cdot$ defined by

$$\begin{array}{rccl} f+g: & \mathbb{R} & \to & \mathbb{R} \\ & t & \mapsto & f(t)+g(t) \end{array} \quad \text{and} \quad \begin{array}{rccl} \alpha \cdot f: & \mathbb{R} & \to & \mathbb{R} \\ & t & \mapsto & \alpha f(t) \end{array}$$

  is a vector space.

Definition 1.2 (*Subspace*)

We say that a non-empty subset S of a vector space V is a subspace *is it is stable by addition and multiplication by a scalar, that is if*

(i) for all $x, y \in S$ we have $x + y \in S$,
(ii) for all $x \in S$ and all $\alpha \in \mathbb{R}$ we have $\alpha x \in S$.

Notice that a subspace is also a vector space!

2 Linear dependency

Definition 2.1 (*Linear combination*)

Let V be a vector space and $A \subset V$. We say that $y \in V$ is a linear combination *of elements of A if there exist $k \in \mathbb{N}$, $x_1, \ldots, x_k \in A$ and $\alpha_1, \ldots, \alpha_k$ such that*

$$y = \sum_{i=1}^{k} \alpha_i x_i.$$

Remember that a linear combination is always a *finite* sum.

Remark 2.1. *If S is a subspace of a vector space V, any linear combination of elements of S belongs to S.*

Definition 2.2 (*Span*)

Let V be a vector space and $A \subset V$. The linear span *of A is the set of all linear combinations of elements of A:*

$$\operatorname{Span}(A) = \left\{ y \,\middle|\, \exists k \in \mathbb{N},\, x_1, \ldots, x_k \in A,\, \alpha_1, \ldots, \alpha_k \in \mathbb{R},\, y = \sum_{i=1}^{k} \alpha_i x_i \right\}.$$

Given vectors $x_1, \ldots x_k \in V$ we will simply write

$$\operatorname{Span}(x_1, \ldots, x_k) = \operatorname{Span}(\{x_1, \ldots, x_k\}) = \left\{ \alpha_1 x_1 + \cdots + \alpha_k x_k \,\middle|\, \alpha_1, \ldots, \alpha_k \in \mathbb{R} \right\}.$$

One can easily verify (exercise!) that $\operatorname{Span}(A)$ is a subspace of V. One can also verify (exercise!) that

$$\operatorname{Span}(A) = \bigcap_{\substack{S \text{ subspace of } V \\ A \subset S}} S,$$

$\operatorname{Span}(A)$ is therefore the **smallest** (for the inclusion $\subset$) **subspace of S that contains** A.

Definition 2.3 (*Linear dependency*)

Vectors $x_1, \dots x_k \in V$ *are* linearly dependent *is there exists* $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ ***that are not all zero*** *such that*

$$\alpha_1 x_1 + \cdots + \alpha_k x_k = 0.$$

They are said to be linearly independent *otherwise.*

Saying that $x_1, \dots, x_k$ are linearly dependent precisely means that one of the vectors $x_1, \dots, x_k$ can be obtained as a linear combination of the others. Indeed if $x_1, \dots, x_k$ are linearly dependent, then we can find $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ that are not all zero (there exists i such that $\alpha_i \neq 0$) such that $\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$. This leads to

$$x_i = \sum_{j \neq i} \frac{-\alpha_j}{\alpha_i} x_j,$$

i.e. the vector x_i can be expressed as a linear combinations of the vectors x_j for $j \neq i$. Conversely if we have for some i, and $\alpha_1, \dots, \alpha_k \in \mathbb{R}$

$$x_i = \alpha_1 x_1 + \cdots + \alpha_{i-1} x_{i-1} + \alpha_{i+1} x_{i+1} + \cdots \alpha_k x_k = 0.$$

then $\alpha_1 x_1 + \cdots + \alpha_{i-1} x_{i-1} - x_i + \alpha_{i+1} x_{i+1} + \cdots \alpha_k x_k = 0$ which gives that $x_1, \dots, x_k$ are linearly dependent.

Theorem 2.1

Let $v_1, \dots, v_n \in V$ *and suppose that we have vectors* $x_1, \dots, x_k \in V$ *such that* $k > n$ *and* $x_i \in \mathrm{Span}(v_1, \dots, v_n)$ *for all* $i \in \{1, \dots, k\}$*. Then* $x_1, \dots, x_k$ *are linearly dependent.*

Theorem 2.1 will be proved in Section 3.

Definition 2.4 (*Basis*)

A family $(x_1, \dots, x_n)$ *of vectors of* V *is a basis of* V *if*

(i) $x_1, \dots, x_n$ *are linearly independent,*

(ii) $\mathrm{Span}(x_1, \dots, x_n) = V$.

Definition 2.5 (*Dimension*)

Let V *be a vector space.*

- *If* V *admits a basis* $(v_1, \dots, v_n)$*, then every basis of* V *has also* n *vectors. We say that* V *has dimension* n *and write* $\dim(V) = n$.
- *Otherwise, we say that* V *has infinite dimension:* $\dim(V) = +\infty$.

The dimension is therefore the minimum number of vector needed to span the vector space. In this course we are going to focus mostly on finite dimensional spaces.

Proof. We proceed by contradiction and assume that there exists two basis $(v_1, \dots, v_n)$ and $(x_1, \dots, x_k)$ of V such that $k \neq n$. Without loss of generality we can assume that $k > n$. For $i = 1, \dots, k$ we have

$$x_i \in V = \mathrm{Span}(v_1, \dots, v_n),$$

because $(v_1, \dots, v_n)$ is a basis of V. We can therefore apply Theorem 2.1 to get that $x_1, \dots, x_{n+1}$ are linearly dependent. This contradicts the fact that $(x_1, \dots, x_k)$ is a basis. □

Proposition 2.1 (***Coordinates***)

Let $(v_1, \ldots, v_n)$ be a basis of V. Then for every $x \in V$ there exists a unique vector $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ such that

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

We say that $(\alpha_1, \ldots, \alpha_n)$ are the coordinates of x in the basis $(v_1, \ldots, v_n)$.

Proof. Existence. $(v_1, \ldots, v_n)$ forms a basis of V therefore $V = \mathrm{Span}(v_1, \ldots, v_n)$. We get that $x \in \mathrm{Span}(v_1, \ldots, v_n)$ which gives that there exists $\alpha_1, \ldots, \alpha_n$ such that $x = \alpha_1 v_1 + \cdot + \alpha_n v_n$.
Uniqueness. Let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{R}$ such that

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n = \beta_1 v_1 + \cdots + \beta_n v_n.$$

This leads to

$$(\alpha_1 - \beta_1) v_1 + \cdots + (\alpha_n - \beta_n) v_n = 0.$$

The vectors $v_1, \ldots, v_n$ are linearly independent because they forms a basis. Consequently $\alpha_1 - \beta_1 = \alpha_2 - \beta_2 = \cdots \alpha_n - \beta_n = 0$, i.e. $(\alpha_1, \ldots, \alpha_n) = (\beta_1, \ldots, \beta_n)$. □

Definition 2.6 (***Lines, hyperplanes***)

Let S be a subspace of $\mathbb{R}^n$.

- *We call S a* line *if* $\dim(S) = 1$.
- *We call S an* hyperplane *if* $\dim(S) = n - 1$.

3 Proof of Theorem 2.1

Notice that it suffices to prove the theorem for $k = n + 1$ because if $x_1, \ldots, x_{n+1}$ are linearly dependent, so are $x_1, \ldots, x_{n+1}, \ldots x_k$. We will therefore show for all $n \geq 1$

$$\begin{aligned} \mathcal{H}_n : \text{« For all } v_1, \ldots, v_n \in V \text{ and all } x_1, \ldots x_{n+1} \in \mathrm{Span}(v_1, \ldots, v_n), \\ \text{the vectors } x_1, \ldots, x_{n+1} \text{ are linearly dependent. »} \end{aligned}$$

Base case: $\mathcal{H}_1$ is true. Indeed, if $x_1, x_2 \in \mathrm{Span}(v_1)$, then there exists $\alpha_1, \alpha_2 \in \mathbb{R}$ such that $x_1 = \alpha_1 v_1$ and $x_2 = \alpha_2 v_1$. If $\alpha_1 = 0$ then $x_1 = 0$ and x_1, x_2 are therefore linearly dependent. Otherwise if $\alpha_1 \neq 0$ then $v_1 = \frac{1}{\alpha_1} x_1$ which then gives $x_2 = \frac{\alpha_2}{\alpha_1} x_1$: x_1, x_2 are linearly dependent.

Induction step: We assume now that $\mathcal{H}_{n-1}$ holds for some $n \geq 2$ and we will show that $\mathcal{H}_n$ holds. We consider therefore $x_1, \ldots, x_{n+1} \in \mathrm{Span}(v_1, \ldots, v_n)$. We can find real numbers $\alpha_{i,j}$ such that

$$\begin{array}{ccccccc} x_1 & = & \alpha_{1,1} v_1 & + & \cdots & + & \alpha_{1,n} v_n \\ x_2 & = & \alpha_{2,1} v_1 & + & \cdots & + & \alpha_{2,n} v_n \\ \vdots & & & & & & \\ x_{n+1} & = & \alpha_{n+1,1} v_1 & + & \cdots & + & \alpha_{n+1,n} v_n. \end{array}$$

We have to show that $x_1, \ldots, x_{n+1}$ are linearly dependent. Let us consider the first line. If $\alpha_{1,1} = \alpha_{1,2} = \cdots = \alpha_{1,n} = 0$, then $x_1 = 0$ which gives then that $x_1, \ldots, x_{n+1}$ are linearly dependent. Otherwise, there exists j such that $\alpha_{1,j} \neq 0$. Without loss of generality we can assume that $\alpha_{1,1} \neq 0$.

$$\begin{array}{ccccccc}
x_1 & = & \alpha_{1,1}v_1 & + & \cdots & + & \alpha_{1,n}v_n \\
x_2 - \frac{\alpha_{2,1}}{\alpha_{1,1}}x_1 & = & 0 & + & \cdots & + & \alpha_{2,n}v_n - \frac{\alpha_{2,1}}{\alpha_{1,1}}\alpha_{1,n}v_n \\
\vdots & & & & & & \\
x_{n+1} - \frac{\alpha_{n+1,1}}{\alpha_{1,1}}x_1 & = & 0 & + & \cdots & + & \alpha_{n+1,n}v_n - \frac{\alpha_{n+1,1}}{\alpha_{1,1}}\alpha_{1,n}v_n.
\end{array}$$

If we define $y_i \stackrel{\text{def}}{=} x_i - \frac{\alpha_{i,1}}{\alpha_{1,1}}x_1$ for $i = 2, \ldots, n+1$ we obtain have $y_i \in \mathrm{Span}(v_2, \ldots, v_n)$. We can now apply the induction hypothesis $\mathcal{H}_{n-1}$ to get that $y_2, \ldots, y_{n+1}$ are linearly dependent. This means that there exists $\beta_2, \ldots \beta_{n+1}$ that are not all zero, such that $\beta_2 y_2 + \cdots + \beta_{n+1}y_{n+1} = 0$ which finally gives

$$\Big(- \beta_2 \frac{\alpha_{2,1}}{\alpha_1, 1} - \cdots - \beta_{n+1}\frac{\alpha_{n+1,1}}{\alpha_{1,1}}\Big)x_1 + \beta_2 x_2 + \cdots + \beta_{n+1}x_{n+1} = 0.$$

Since $\beta_2, \ldots, \beta_{n+1}$ are not all zero we get that $x_1, \ldots, x_{n+1}$ are linearly dependent. $\mathcal{H}_n$ is proved.