Vector Space

Definition (Vector Space). A set $\mathcal{V}$ is called a $\mathbb{R}$ vector space, if it is equipped with an addition operation $+: \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$ and a scalar multiplication operation $\cdot: \mathbb{R} \times \mathcal{V} \rightarrow \mathcal{V}$ such that the following properties are satisfied for all $a, b, c \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{R}$.

i. $a + b = b + a$ and $a + (b + c) = (a + b) + c$
ii. There is an element $0 \in V$ such that $0 + a = a + 0 = a$ [*Neutral element*]
iii. There is an element called $-a$ such that $a + (-a) = (-a) + a = 0$
iv. $\alpha(a + b) = \alpha a + \alpha b$ and $(\alpha + \beta)a = \alpha a + \beta a$ [*Distributive property*]
v. $\alpha(\beta a) = (\alpha\beta)a$ [*Associative property*]

The elements $a \in \mathcal{V}$ are called *vectors.*

Remark

A "vector" multiplication ab with $a, b \in \mathbb{R}^n$ is not defined. Theoretically, we could define an element-wise multiplication, such that $c = ab$ with $c_j = a_j b_j$. This "array multiplication" is common to many programming languages but makes mathematically limited sense using the standard rules for matrix multiplication, since the dimensions of the vectors do not match.

Only the following multiplications for vectors are defined: $ab^{\mathrm{T}} \in \mathbb{R}^{n \times n}$ (outer product), $a^{\mathrm{T}}b \in \mathbb{R}$ (inner/scalar/dot product).

There are many different vector spaces that will pop up in data science:

a. Most important: $\mathbb{R}^n$, the $n - dimensional\ real\ vector\ space$ that consists of tuples of n real numbers. The space $\mathbb{C}^n$ of complex vectors is less prominent but may occur as well.

b. As important: *Spaces of functions!* The set of all functions $f: A \rightarrow \mathbb{R}$ from some set A into the real numbers form a vector space as well. Such a space may be used to model a set of decision functions that we want to build to predict an output y for given data $x \in A$. If we want to predict more than just one number for a data point x we consider the space of functions $f: A \rightarrow \mathbb{R}^n$, and those functions form a vector space as well.

A. Inner Product

(Definition) Let V be a real vector space. An inner product on V is a map $\langle \cdot\, ,\ \cdot \rangle : V \times V \rightarrow \mathbb{R}$ with the following properties:

i. For all $x, y, z \in V$ and $\alpha \in \mathbb{R}$, we have linearity in the first argument, i.e.,

$$\langle x + \alpha y, z \rangle = \langle x, z \rangle + \alpha \langle y, z \rangle$$

ii. For all $x, y \in V$, we have symmetry i.e.,

$$\langle x, y \rangle = \langle y, x \rangle$$

iii. For all $x \in V$, we have positive definiteness i.e.,

$$\langle x, x \rangle \geq 0 \; and \; \langle x, x \rangle = 0 \; only \; if \; x = 0$$

(Remark)

1. Properties (i) and (ii) together imply that $\langle \cdot, \ \cdot \rangle$ is also linear in the second argument. Hence, an inner product is a so-called bilinear map.
2. For vector spaces V over $\mathbb{C}$, we require ***sesquilinearity***, i.e.,
$$\langle x, y \rangle = \overline{\langle y, x \rangle}$$
and
$$\langle \lambda_1 a + b, \lambda_2 c + d \rangle = \overline{\lambda_1} \langle a, d \rangle + \overline{\lambda_1} \lambda_2 \langle a, c \rangle + \lambda_2 \langle b, c \rangle + \langle b, d \rangle$$

for $\lambda_1, \lambda_2 \in \mathbb{C}$ and $a, b, c, d \in V$.

(Example)

1. On $\mathbb{R}^n$, there is the standard inner product, also called dot product $\langle x, y \rangle_{\mathbb{R}^n} := \sum_{i=1}^{n} x_i y_i$. The inner product can also be written using matrix-vector multiplication as $x^T y$.
2. On the space of functions $u : \ \Omega \to \mathbb{R}$, we can define the so-called L^2-inner product
$$\langle u, v \rangle_{L^2} = \int_{\Omega} u(x) v(x) dx$$
3. In the spirit of the first point, we can define for any matrix $A \in \mathbb{R}^{n \times n}$ a bilinear map
$$\langle x, y \rangle_A := y^T A x$$
This map is symmetric when A is so and is an inner product when A is positive definite, i.e., when $x^T A x > 0$ for $x \neq 0$.

It is of fundamental importance that *inner products induce norms*, i.e., for any inner product we can define

$$\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$

(Theorem) Cauchy – Schwartz Inequality for Inner Product Spaces

Let V be a $\mathbb{R}$ vector space with inner product $\langle \cdot \ , \ \cdot \rangle$. Then, for all $x, y \in V$, we have

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle$$

Proof.

Without loss of generality we may assume, $\langle x, y \rangle \neq 0$. Then, for $\lambda \neq 0$, we have:

$$0 \leq \langle \lambda x + y, \lambda x + y \rangle = \lambda^2 \langle x, x \rangle + 2\lambda \langle x, y \rangle + \langle y, y \rangle$$

For any $\lambda \in \mathbb{R}$, the expression $\langle \lambda x + y, \lambda x + y \rangle$ represents the inner product of a vector with itself, which is always non-negative.

Now, since $\langle \lambda x + y, \lambda x + y \rangle \geq 0$, the discriminant of this quadratic must be less than or equal to zero (for it to not have any real roots or to have a double root). Therefore,

$$(2\langle x, y \rangle)^2 - 4 \cdot \langle x, x \rangle \cdot \langle y, y \rangle \leq 0$$

This gives us the required expression:

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle$$

B. Metric

(Definition) Let X be a set. A metric on X is a map

$$d: X \times X \to (0, \infty)$$

with the following properties:

i. For all $x, y \in X$, it holds that $d(x, y) = d(y, x)$ [*symmetry*]
ii. It holds that $d(x, y) \geq 0, = 0$ exactly if $x = y$ [*positive definite*]
iii. For all $x, y, z \in X$, it holds that all $d(x, z) \leq d(x, y) + d(y, z)$ [*triangle inequality*]

Similar to the case of norm, we can define convergence of sequences in metric spaces. We say $x_n \to x$ with respect to the metric d if $d(x_n, x) \to 0$.

(Example)

Let X be a unit circle in $\mathbb{R}^2$, i.e., the set $X = \{x \in \mathbb{R}^2 \mid \|x\|_2 = 1\}$.

Since X is a subset of the vector space $\mathbb{R}^2$, we could measure the distance of $x, y \in X$ by $\|x - y\|_2$ which would be the distance we measure with a ruler. Alternatively, we could also measure the distance by the angle between x, y.

More generally, the angel between two vectors in $\mathbb{R}^d$ is given by

$$\measuredangle(x, y) := arccos\left(\frac{\langle x, y \rangle}{\|x\|\|y\|}\right)$$

And this angle does indeed define a metric on the unit sphere $S^{d-1} = \{x \in \mathbb{R}^d | \|x\|_2 = 1\}$.

(Remarks)

Similar to how inner product induces norms, it is also true that norms induce metrics: if the set X is also a vector space and $\|\cdot\|$ is a norm on this space, then:

$$d(x, y) := \|x - y\|$$

is in fact a metric.

Metrics are often used when the underlying space is not a vector space.

Now, we have:

$$inner\ product \xrightarrow{induces} norm \xrightarrow{induces} metric$$

which means, the metric is the most general notation. Note that a metric need not be induced by a norm. Consider, e.g., the discrete metric:

$$d(x,y) = \begin{cases} 0, & x = y \\ 1, & x \neq y \end{cases} \quad x, y \in \mathbb{R}$$

Clearly there are $\alpha \in \mathbb{R}$, for which $d(\alpha x, 0) \neq |\alpha| d(x, 0)$ with $x \neq 0$, i.e., the homogeneity property fails.

C. Metric Space

A metric space is a set equipped with a metric. More formally, a metric space is a pair (X, d), where:

- X is a set (the elements of which can be points, vectors, or other objects)
- d is a metric on X, i.e., a function that satisfies the above properties of a metric

Thus, a metric space is the complete structure that includes both the set X and the distance function d.

(Example)

Consider the set $\mathbb{R}$ and the standard absolute value metric $d(x, y) = |x - y|$.

- The metric here is $d(x, y) = |x - y|$, which tells us the distance between any two real numbers x and y.
- The metric space is the pair $(\mathbb{R}, d)$, where $\mathbb{R}$ is the set of real numbers, and $d(x, y) = |x - y|$ is the metric.

(Remark)

Let X be a set equipped with a metric d.

1. For $x \in X$ and $\epsilon > 0$ we define
$$B_\epsilon(x) := \{y \in X \mid d(x, y) < \epsilon\}$$
and call $B_\epsilon(x)$ the *open ball* of radius ϵ around x or the ϵ-neighborhood of x.

2. A set $U \subset X$ is called *open*, if for all $x \in U$ there exists an $\epsilon > 0$ (which is allowed to depend on x) such that $B_\epsilon(x) \subset U$

3. A set $A \subset X$ is called *closed*, if its complement set $A^c = X \setminus \{A\}$ is open.

(Remark)

Let (X, d) be a metric space.

1. If some given subset $U \subset X$ is open or not depends on the set X which contains U.
2. There are sets which are neither open, nor closed – think of the half-open interval $(a, b]$ as subset of ℝ. On the other hand, there are sets which are both open and closed, called "*clopen*". For instance, both the set X of the original metric space and the empty set ϕ are both open and closed.

(Remark)

In a metric space X, a set A is closed if and only if the limit $x \in X$ of every convergent sequence $x_n \xrightarrow{n \to \infty} x$ with $(x_n)_{n \in \mathbb{N}} \subseteq A$, one has $x \in A$. Put differently, convergent sequences in A cannot leave the set.

(Example)

We consider the real line $X = \mathbb{R}$ with the metric $d(x, y) = |x - y|$.

- The intervals (a, b) are open. (Given some $x \in (a, b)$, can you find an $\epsilon > 0$ such that $B_\epsilon(x) = (x - \epsilon, x + \epsilon) \subset (a, b)$?)
- The intervals $[a, b]$ are closed. (Which follows from the fact that if $x_n \leq a$ and $x_n \to x$, then $x \leq a$).
- The intervals $]a, b]$ and $[a, b[$ are neither closed, nor open, while ℝ and the empty set ϕ are both open and closed.

Foundation of Hilbert Space

Key Features of Hilbert Space

The key features of a Hilbert space are:

1. **Vector space**: It is a vector space, meaning it has elements (vectors) that can be added together and scaled by numbers (scalars), following certain rules.

2. **Inner product**: Each Hilbert space has an inner product, which is a function that takes two vectors and returns a scalar (usually a real or complex number). This inner product provides a notion of angle and length, allowing concepts like orthogonality and distance between vectors.
3. **Completeness**: Hilbert spaces are complete with respect to the norm induced by the inner product. This means that every Cauchy sequence (a sequence where the vectors get arbitrarily close to each other) in the Hilbert space has a limit that also lies within the space.

4. **Infinite dimensions**: While Hilbert spaces can be finite-dimensional (like regular Euclidean space), they are most often used in contexts where the space is infinite-dimensional, such as in quantum mechanics or the theory of partial differential equations.

Applications of Hilbert Space

- **Quantum mechanics**: In quantum theory, the state of a physical system is represented as a vector in a Hilbert space. Observables (like position, momentum, and energy) are represented by operators acting on this space.
- **Fourier analysis**: Hilbert spaces are used to represent functions as sums of basic waveforms, which is a key idea in Fourier series and transforms.
- **Signal processing**: Functions or signals can be represented in Hilbert spaces, where operations like filtering and signal reconstruction are defined in terms of the inner product.