COMP0014 Tutorial 1

Changmin Yu

Basic Linear Algebra

Vector Space

A vector space over a field (e.g., $\mathbb{R}$ or $\mathbb{C}$) is a set V of elements, or 'vectors', together with two binary operators.

- *vector addition* denoted for $v_1, v_2, \in V$ by $v_1 + v_2$, where $v_1 + v_2 \in V$, i.e., a vector space is closed under addition.
- *scalar multiplication* denoted for $\lambda \in \mathbb{R}$ and $v \in V$ by λv, where $\lambda v \in V$, so that the vector space is closed under scalar multiplication.

Vector spaces satisfy the following 8 rules:

- Addition is commutative, i.e. for all $v_1, v_2 \in V$
$$v_1 + v_2 = v_2 + v_1 \tag{1}$$
- Addition is associative, i.e. for all $v_1, v_2, v_3 \in V$
$$v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3 \tag{2}$$
- There exists a unique element $0 \in V$, called the null or zero vector, such that for all $v \in V$
$$v + 0 = v \tag{3}$$
- For all $v \in V$ there exists an additive negative or inverse vector $v' \in V$ such that
$$v + v' = 0 \tag{4}$$
- Scalar multiplication is distributive over scalar addition, i.e. for all $\lambda, \mu \in \mathbb{R}$, and $v \in V$
$$(\lambda + \mu)v = \lambda v + \mu v \tag{5}$$
- Scalar multiplication is distributive over vector addition, i.e. for all $\lambda \in \mathbb{R}$ and $v_1, v_2 \in V$
$$\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2 \tag{6}$$
- Scalar multiplication of vectors is 'associative', i.e. for all $\lambda, \mu \in \mathbb{R}$ and $v \in V$
$$\lambda(\mu v) = (\lambda\mu)v \tag{7}$$
- Scalar multiplication has an identity element, i.e. for all $v \in V$
$$1 \cdot v = v \tag{8}$$
where 1 is the multiplicative identity in $\mathbb{R}$.

Spanning Sets, Linear Independence, Bases

First consider 2-dimensional space, $\mathbb{R}^2$, an origin O, and two non-zero and non-parallel vectors v_1 and v_2. Then any vector $v \in \mathbb{R}^2$, we have

$$v = \lambda v_1 + \mu v_2 \tag{9}$$

for scalars $\lambda, \mu \in \mathbb{R}$. We say that the set $\{v_1, v_2\}$ **spans** the set of vectors lying in $\mathbb{R}^2$.

Definition: Spanning set. We say that $S = \{v_1, \ldots, v_n\}$ spans a vector space V if for all $v \in V$, v can be expressed as a linear combination of the vectors in S, i.e. for all $v \in V$

$$v = \sum_{i=1}^{n} \lambda_i v_i \tag{10}$$

where $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. In such cases, we say that S spans V.

Definition: Linear independence A set of vectors $S = \{v_1, \ldots, v_n\}$ is said to be a linearly independent set if

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \quad \Rightarrow \quad \lambda_i = 0, \quad i = 1, \ldots, n \tag{11}$$

Definition: Basis We say that the set $S = \{v_1, \ldots, v_n\}$ is a *basis* for a vector space if S is a spanning set and S is linearly independent.

Affine Spaces

Affine spaces are very similar to vector spaces (AKA *linear spaces*) except that they contain a translation. Like vector spaces, affine spaces have the shape of a linear "object" – a point, line, plane, or hyperplane. However, unlike vector spaces, affine spaces do not necessarily need to include the special zero vector ($0 \in V$) – see rule 3 of vector spaces.

Another way of thinking of affine spaces is that all affine spaces correspond to or are equivalent to a vector space (you can create a bijection between the elements of a vector space and the elements of an affine space), but they are translated so that they do not necessarily contain the origin (0). That is, they are vector spaces that have been translated in space.

We will talk more about affine spaces and affine transformations when we discuss matrices next week.

Norms, magnitude, and inner products

A **norm** is a function that maps vectors to the set of non-negative real numbers and obeys certain properties. More formally, we can say a norm is a function $f : V \to \mathbb{R}$, where V is a vector space. There are multiple norms that exist, and several of these norms are important in the field of machine learning, but for now we will focus on the most common norm: the L2 norm or *Euclidean norm*:

$$\|v\| = \|v\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2} \tag{12}$$

The **magnitude** of a mathematical object refers to the size of that object. Magnitudes are defined for multiple types of objects other than vectors, but in the case of vectors, the magnitude is just the Euclidean norm, so magnitude and norm are synonyms. They both refer to the length of a vector in Euclidean space, which is calculated by the equation above: the square root of the sum of squares.

Norms obey the following properties:

- Subadditivity, or triangle inequality, i.e. if f is a norm, then for all $v_1, v_2 \in V$

$$f(v_1 + v_2) \leq f(v_1) + f(v_2) \tag{13}$$

- Scalability, i.e. for all $v \in V$ and $\lambda \in \mathbb{R}$

$$f(\lambda v) = \lambda f(v) \tag{14}$$

- The zero vector is (the only vector) assigned a norm of 0.

$$f(\mathbf{0}) = 0 \tag{15}$$

$$f(v) = 0 \implies v = \mathbf{0} \tag{16}$$

In contrast to norms, an **inner product** takes TWO vectors in a vector space and maps them to a scalar (not necessarily non-negative), obeying a different set of properties. More formally, we can say an inner product is a function $f : V \times V \to \mathbb{R}$ (or $\mathbb{C}$, but let's stick to $\mathbb{R}$ for now). Like norms, different inner products can be defined. However, they generally represent a way of numerically indicating "similarity" between two vectors in a vector space. In particular, the standard inner product for Euclidean space is the *dot product* or *scalar product*, which for two vectors $v_1 = [v_1^{(1)}, ..., v_1^{(n)}]$ and $v_2 = [v_2^{(1)}, ..., v_2^{(n)}]$ is defined as:

$$v_1 \cdot v_2 = \sum_{i=1}^{n} v_1^{(i)} v_2^{(i)} \tag{17}$$

Therefore:

$$\|v\| = \sqrt{v \cdot v} \tag{18}$$

Inner products using scalars in $\mathbb{R}$ obey the following properties:

- Symmetry, i.e. if f is an inner product, then for all $v_1, v_2 \in V$

$$f(v_1, v_2) = f(v_2, v_1) \tag{19}$$

- Linearity, i.e. for all $v_1, v_2, v_3 \in V$ and $\lambda_1, \lambda_2 \in \mathbb{R}$

$$f(\lambda_1 v_1 + \lambda_2 v_2, v_3) = \lambda_1 f(v_1, v_3) + \lambda_2 f(v_2, v_3) \tag{20}$$

- Positive-definiteness, i.e. if x is not 0, then

$$f(x, x) > 0 \tag{21}$$