CSE 840: Computational Foundations of Artificial Intelligence Aug 28, 2023

Vector Spaces, Basis and Dimension, Direct Sum

Instructor: Vishnu Boddeti *Scribe: Batuhan Yucer, Onur Can Yucedag*

1 Vector Spaces

Definition 1 *A **group** is a set of elements* $\mathbf{G}$ *with an operation* $+ : \mathbf{G} \times \mathbf{G} \to \mathbf{G}$ *that satisfies the following properties:*
(+ is used as a representative symbol for any operation, and does not necessarily mean "addition")

(P1) ***Associativity:*** $\forall a, b, c \in \mathbf{G} : (a + b) + c = a + (b + c)$

(P2) ***Identity element:*** $\exists e \in \mathbf{G}, \forall g \in \mathbf{G} : e + g = g + e = g$
E.g. *For the addition operation, the identity element is* 0.

(P3) ***Inverse element:*** $\forall a \in \mathbf{G}, \exists b \in \mathbf{G} : a + b = b + a = e$
E.g. *For the addition operation,* $b = -a$.

*A group is called a **commutative group** (or **abelian group**) if it also satisfies the following property:*

(P4) ***Commutativity:*** $\forall a, b \in \mathbf{G} : a + b = b + a$

Examples:

- $(\mathbb{R}^n, +)$ is a group.
- $(\mathbb{R}^+, \cdot)$ is a group.
- $(\mathbb{R}^-, \cdot)$ is **not** a group.
- Given the set of permutation matrices

  $\mathbf{S_n} := \left\{\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\} \;\middle|\; \pi \text{ is bijective}\right\}$

  and the combination operation

  $\circ : \mathbf{S_n} \times \mathbf{S_n} \to \mathbf{S_n}$

  $\pi_1 \circ \pi_2(i) := \pi_1(\pi_2(i))$

  $(\mathbf{S_n}, \circ)$ is a group.

Definition 2 *A **field** is a set of elements* $\mathbf{F}$ *with two operations* $+, \cdot : \mathbf{F} \times \mathbf{F} \to \mathbf{F}$ *that satisfies the following properties:*

(P1) $(\mathbf{F}, +)$ *is a commutative group with the identity element* 0*.*

(P2) $(\mathbf{F} \setminus \{0\}, \cdot)$ *is a commutative group with the identity element* 1*.*

(P3) ***Distributivity:*** $\forall a, b, c \in \mathbf{F} : a \cdot (b + c) = a \cdot b + a \cdot c$

Examples:

- $(\mathbb{R}, +, \cdot)$ is a field.
- $(\mathbb{C}, +, \cdot)$ is a field.
- Given $n \in \mathbb{Z}$, the set

  $$\mathbf{Z_n} := \big\{0, 1, \ldots, n-1\big\}$$

  and the operations

  $$a +_n b := (a + b) \mod n$$
  $$a \cdot_n b := (a \cdot b) \mod n$$

  (subscript n denotes that the operations are defined on these n integers)

  $(\mathbf{Z_n}, +_n, \cdot_n)$ is a field **if and only if** n is a prime.

Definition 3 *Let* $\mathbf{F}$ *be a field with identity elements* 0 *and* 1*. A **vector space** defined over the field* $\mathbf{F}$ *is a set* $\mathbf{V}$ *with two mappings*

$+ : \mathbf{V} \times \mathbf{V} \to \mathbf{V}$ *(vector addition)*
$\cdot : \mathbf{F} \times \mathbf{V} \to \mathbf{V}$ *(scalar multiplication)*

that satisfies the following properties:

(P1) $(\mathbf{V}, +)$ *is a commutative group.*

(P2) ***Multiplicative identity:*** $\forall v \in \mathbf{V} : 1 \cdot v = v$

(P3) ***Distributivity:*** $\forall a, b \in \mathbf{F}, \forall u, v \in \mathbf{V} :$

$$a \cdot (u + v) = a \cdot u + a \cdot v$$
$$(a + b) \cdot u = a \cdot u + b \cdot u$$

Remark 4 *The elements of* $\mathbf{V}$ *are called **vectors**, and the elements of* $\mathbf{F}$ *are called **scalars**.*

Examples:

- $\mathbb{R}^n$ with the standard operations $(+, \cdot)$ is a vector space.
- **Function spaces:**
  $\mathcal{F}(X, \mathbb{R}) := \{f : X \to \mathbb{R}\}$
  is the space of all real-valued function on a set X. Given the operation
  $+ : \mathcal{F}(X, \mathbb{R}) \times \mathcal{F}(X, \mathbb{R}) \to \mathcal{F}(X, \mathbb{R})$
  $(f + g)(x) := f(x) + g(x)$
  where $f, g \in \mathcal{F}(X, \mathbb{R})$ and $x \in X$
  and the operation
  $\cdot : \mathbb{R} \times \mathcal{F}(X, \mathbb{R}) \to \mathcal{F}(X, \mathbb{R})$
  $(\lambda \cdot f)(x) := \lambda \cdot f(x)$
  where $\lambda \in \mathbb{R}$, $f \in \mathcal{F}(X, \mathbb{R})$ and $x \in X$
  $(\mathcal{F}(X, \mathbb{R}), +, \cdot)$ is a real-vector space.
- $\mathcal{C}(X) := \{f : X \to \mathbb{R} \mid f \text{ is continuous}\}$ also forms a vector space.
- $\mathcal{C}^r([a, b]) := \{f : [a, b] \to \mathbb{R} \mid f \text{ is } r \text{ times continuously differentiable}\}$ also forms a vector space.

Definition 5 *Let* $\mathbf{V}$ *be a vector space over field* $\mathbf{F}$*, and* $\mathbf{U} \subset \mathbf{V}$ *be a non-empty set.* $\mathbf{U}$ *is a* ***subspace*** *of* $\mathbf{V}$ *if it is* *closed* *under linear combinations:*

$\forall \lambda, \mu \in \mathbf{F},\ \forall u, v \in \mathbf{U} : \lambda \cdot u + \mu \cdot v \in \mathbf{U}$

Examples:

- $\mathcal{C}(X)$ is a subspace of $\mathcal{F}(X, \mathbb{R})$.
- The sets of symmetric matrices of size $n \times n$ is a subspace of $\mathbb{R}^{n \times n}$ (all matrices of size $n \times n$; real valued matrices are assumed).
- The set $\{u, v\} \subset \mathbf{V}$ is **not** a subspace of $\mathbf{V}$; because $\lambda \cdot u + \mu \cdot v \notin \{u, v\}$ for an arbitrarily chosen $\lambda, \mu \in \mathbb{R}$, meaning that the set is not closed under linear combinations.

Definition 6 *Let* $\mathbf{V}$ *be a vector space over field* $\mathbf{F}$*. Given* $u_1, u_2, \ldots, u_n \in \mathbf{V}$ *and* $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbf{F}$*,* $\sum_{i=1}^{n} \lambda_i u_i$ *is called a linear combination.*

The set of all existing linear combinations of $(u_1, u_2, \ldots, u_n)$ *is the* ***span*** *(or* ***linear hull****) of* $(u_1, u_2, \ldots, u_n)$*:*

span$(u_1, u_2, \ldots, u_n) := \left\{ \sum_{i=1}^{n} \lambda_i u_i \;\middle|\; \lambda_i \in \mathbf{F} \right\}$

The set $\mathbf{U} = \{u_1, u_2, \ldots, u_n\}$ *is the* ***generator*** *of* ***span***$(\mathbf{U})$*.*

Definition 7 *A set of vectors* $v_1, v_2, \ldots, v_n$ *are* ***linearly independent*** *if the following condition holds:*

$\sum_{i=1}^{n} \lambda_i v_i \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$

Examples:

- Vectors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3$ are linearly independent.
- Functions $\sin(x)$ and $\cos(x)$ are linearly independent.
- Any set of $d+1$ vectors in $\mathbb{R}^d$ are linearly dependent.

2 Basis and Dimension

Definition 8 *A subset* $\mathbf{B}$ *of a vector space* $\mathbf{V}$ *is called a (Hamel)* ***basis*** *if the following conditions hold:*

(P1) *span*$(\mathbf{B}) = \mathbf{V}$

(P2) $\mathbf{B}$ *is linearly independent.*

The ***Hamel basis*** *concerns only finite linear combinations, even for infinite dimensional vector spaces.*

Examples:

- The canonical basis of $\mathbb{R}^3$
$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$
- Another basis of $\mathbb{R}^3$:
$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \text{ or } \begin{pmatrix} 0.5 \\ 0.8 \\ 0.4 \end{pmatrix}, \begin{pmatrix} 1.8 \\ 0.3 \\ 0.3 \end{pmatrix}, \begin{pmatrix} -2.2 \\ -1.3 \\ 3.5 \end{pmatrix}$$

Proposition 9 *If $\mathbf{U} = \{u_1, ..., u_n\}$ spans a vector space $\mathbf{V}$, then the set $\mathbf{U}$ can be reduced to a basis of $\mathbf{V}$*

Proof:

1. If $\mathbf{U}$ is already linearly independent, then $\mathbf{U} = \{\mathbf{u_1}, ..., \mathbf{u_n}\}$ in $\mathbf{V}$ is a basis of $\mathbf{V}$.
2. If vectors in $\mathbf{U}$ are linearly dependent: $\exists a \in \mathbf{U}$ that is a linear combination of the other vectors in $\mathbf{U}$. Remove vector a. Repeat this step until remaining vectors are linearly independent.

□

Definition 10 *A vector space is called finite-dimensional if it has at least one finite basis.*

Proposition 11 *Let $\mathbf{U} = \{u_1, ..., u_n\} \subset \mathbf{V}$ be a set of linearly independent vectors, and let $\mathbf{V}$ be a finite-dimensional vector space, then $\mathbf{U}$ can be extended to a basis of $\mathbf{V}$.*

Proof: Let $\{w_1, w_2, ..., w_m\}$ be a basis of $\mathbf{V}$. Consider the set $\{u_1, u_2, ..., u_n, w_1, w_2, ..., w_m\}$.

Remove vectors "from the end" until the remaining vectors are linearly independent

- Remaining set spans $\mathbf{V}$
- Remaining set is linearly independent by construction
- Remaining set contains $\mathbf{U}$

As a consequence of this proof, we can extend every independent set to a basis.

This proof only works on finite-dimensional vector spaces. For infinite-dimensional spaces, Zorn's Lemma is needed to prove the proposition.

□

Corollary 12 *Let* $\mathbf{V}$ *be a finite-dimensional vector space, then any two bases of* $\mathbf{V}$ *have the same length.*

Definition 13 *The length of a basis of a finite dimensional vector space is called the* **dimension** *of* $\mathbf{V}$*.*

3 Sum and Direct Sum

Definition 14 *Assume that* $\mathbf{U_1}$*,* $\mathbf{U_2}$ *are subspaces of* $\mathbf{V}$*. The* ***sum*** *of the two spaces is defined as:*

$$\mathbf{U_1} + \mathbf{U_2} := \{u_1 + u_2 | u_1 \in \mathbf{U_1}, u_2 \in \mathbf{U_2}\}$$

The sum is called a ***direct sum****, if each element in the sum can be written in exactly one way:*

$$\mathbf{U_1} \bigoplus \mathbf{U_2}$$

Proposition 15 *Suppose* $\mathbf{V}$ *is a finite-dimensional vector space, and* $\mathbf{U} \subset \mathbf{V}$ *is a subspace. Then there exists a subspace* $\mathbf{W} \subset \mathbf{V}$*, such that* $\mathbf{U} \bigoplus \mathbf{W} = \mathbf{V}$*.*

Proof: Let the set $\{u_1, u_2, ..., u_n\}$ be a basis of $\mathbf{U}$. Extend it to a basis of $\mathbf{V}$, say the resulting set is:

$$\{\underbrace{u_1, u_2, ..., u_n}_{\text{span}(\mathbf{U})}, \underbrace{v_1, v_2, ..., v_m}_{\text{span}(\mathbf{W})}\}$$

$$\mathbf{W} = span\{v_1, v_2, ..., v_m\}$$

For non-overlapping subspaces, the direct sum helps to define the notion of space and its complement. Here, the subspace $\mathbf{W}$ is the complement of the subspace $\mathbf{U}$; together, they form the vector space $\mathbf{V}$.

□