CSE 840: Computational Foundations of Artificial Intelligence January 15, 2024

Vector Spaces, Basis, and Dimension, Direct Sum

Instructor: Vishnu Boddeti *Scribe: Andrew John J, Laxmi Vatsalya Daita*

Vector Spaces

First Principles, Groups $\mathcal{G}$ then Field.

Definition 1. ***A*** *set G of elements with an operation $+ : G \times G \to G$ is called a group if the following properties hold:*

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

(P2) Identity element: $\exists e \in G, \forall g \in G : e + g = g + e = g$

(P3) Inverse element: $\forall a \in G, \exists b \in G : a + b = b + a = e$

The group is called a commutative group (Abelian group) if we have an additional property:

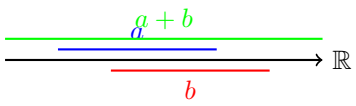
(P4) $\forall a, b \in G : a + b = b + a$

Examples

- $(\mathbb{R}^n, +)$ is a group
- $(\mathbb{R}^*, \cdot)$ is a group
- $(\mathbb{R}^-, \cdot)$ is not a group
- $S_n := \{\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\} \mid \pi \text{ is bijective}\}$
  $\circ : S_n \times S_n \to S_n$
  $\pi_1 \circ \pi_2(i) = \pi_1(\pi_2(i))$
  $(S_n, \circ)$ is a group.
- The rotations by $0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ or vertices of a square form a group.

Visual Representation of a Group

Explanation: For example, $(\mathbb{R}, +)$ is a group because the real numbers are closed under addition, addition is associative, 0 is the identity element, and every number has an additive inverse.



Definition 2. ***A*** *set F with two operations $+, \cdot : F \times F \to F$ is called a field if the following properties hold:*

(P1) $(F, +)$ is a commutative group with identity element 0

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

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

Examples

- $(\mathbb{R}, +, \cdot)$
- $(\mathbb{C}, +, \cdot)$
- $n \in \mathbb{Z}$, consider $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$
  $a +_n b := (a + b) \bmod n$
  $a \cdot_n b := (a \cdot b) \bmod n$
  Then $(\mathbb{Z}_n, +_n, \cdot_n)$ is a field if and only if n is prime

Definition 3. ***L****et $n \in \mathbb{Z}$ and define $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ with the operations:*

- *Addition modulo n:* $a +_n b := (a + b) \mod n$
- *Multiplication modulo n:* $a \cdot_n b := (a \cdot b) \mod n$

Then, $(\mathbb{Z}_n, +_n, \cdot_n)$ forms a field if and only if n is prime.

Proof. To verify that $(\mathbb{Z}_n, +_n, \cdot_n)$ is a field, the following must hold:

(F1) $(\mathbb{Z}_n, +_n)$ is an abelian group.

(F2) $(\mathbb{Z}_n \setminus \{0\}, \cdot_n)$ is an abelian group.

(F3) Multiplication distributes over addition.

- **(F1) Additive Group:** For any n, $(\mathbb{Z}_n, +_n)$ forms a cyclic group under addition modulo n, which is commutative.
- **(F2) Multiplicative Group:** For $(\mathbb{Z}_n \setminus \{0\}, \cdot_n)$ to be a group, every nonzero element of $\mathbb{Z}_n$ must have a multiplicative inverse. This holds if and only if n is prime because:
  - When n is prime, all integers $1, 2, \ldots, n-1$ are coprime to n, and each has a unique multiplicative inverse modulo n.
  - When n is not prime, there exist elements (e.g., divisors of n) in $\mathbb{Z}_n \setminus \{0\}$ that do not have an inverse, violating the field property.
- **(F3) Distributive Property:** The distributive property of multiplication over addition holds for any n due to the modular arithmetic definition.

Conclusion: For $(\mathbb{Z}_n, +_n, \cdot_n)$ to be a field, n must be prime. □

Remark: If n is composite, $(\mathbb{Z}_n, +_n, \cdot_n)$ forms a commutative ring with unity but not a field.

Definition 4. *Let F be a field with identity elements 0 and 1. A vector space defined over the field F is a set V with a mapping $+ : V \times V \to V$ ("vector addition") and $\cdot : F \times V \to V$ ("scalar multiplication") such that:*

(P1) $(V, +)$ is a commutative group

(P2) Multiplicative identity: $\forall v \in V : 1 \cdot v = v$

(P3) Distributive property: $\forall a, b \in F, u, v \in V$
$a \cdot (u + v) = a \cdot u + a \cdot v$
$(a + b) \cdot u = a \cdot u + b \cdot u$

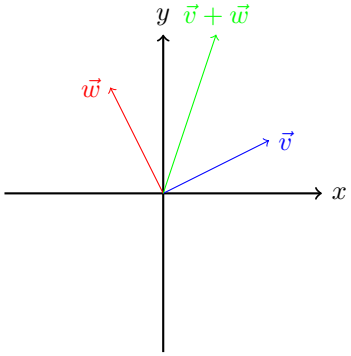
Remark: Elements of V are called vectors, elements of field F are scalars.

Examples

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

Visual Representation of a Vector Space

Explanation: For example, $\mathbb{R}^2$ is a vector space with the standard operations of addition and scalar multiplication.



Subspaces and Linear Combinations

In data science, a key goal is to find simplified representations of complex objects. Often, these objects are represented as m-tuples, v_m, with large m. Assuming addition and scaling are well-defined, these objects form a vector space, typically over the field of real numbers, equipped with Euclidean distance $\mathbb{E}_m$. Examples include medical recordings (e.g., electroencephalograms, electrocardiograms), sound data, or images, where m can reach millions. A fundamental question arises: are all m parameters necessary to describe these objects, or can a smaller set of parameters suffice? This leads to the concept of a vector subspace.

Definition 5. ***L****et V be a vector space, $U \subset V$ non-empty set. We call U a subspace of V if it is closed under linear combinations:*

$\forall \lambda, \mu \in F, \forall u, v \in U : \lambda u + \mu v \in U$

Vector subspaces arise in the decomposition of a vector space. The converse, composition of vector spaces $\mathcal{U} = (U, S, +, \cdot)$, $\mathcal{V} = (V, S, +, \cdot)$ is also defined in terms of linear combination. A vector $\mathbf{x} \in \mathbb{R}^3$ can be obtained as the linear combination

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ x_3 \end{bmatrix},$$

but also as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - a \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a \\ x_3 \end{bmatrix},$$

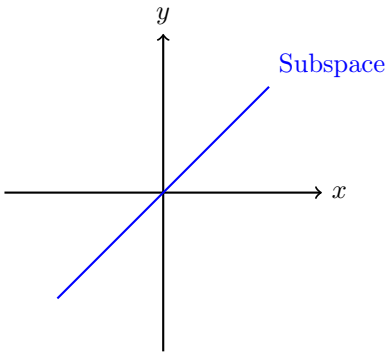
for some arbitrary $a \in \mathbb{R}$. In the first case, $\mathbf{x}$ is obtained as a unique linear combination of a vector from the set $U = \left\{ \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} \middle| x_1 \in \mathbb{R} \right\}$ with a vector from $V = \left\{ \begin{bmatrix} 0 \\ x_2 \\ x_3 \end{bmatrix} \middle| x_2, x_3 \in \mathbb{R} \right\}$. In the second case, there is an infinity of linear combinations of a vector from $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} \middle| x_1, x_2 \in \mathbb{R} \right\}$ to the vector $\mathbf{x}$. This is captured by a pair of definitions to describe vector space composition.

Examples

- $\mathcal{C}(X)$ is a subspace of $\mathcal{F}(X, \mathbb{R})$
- The set of symmetric matrices of size $n \times n$ is a subspace of $\mathbb{R}^{n \times n}$
- Consider set $\{u, v\} \rightarrow$ not a subspace. $\lambda u + \mu v \notin \{u, v\}$ for $\lambda, \mu \in \mathbb{R}$

Visual Representation of a Subspace

Explanation: A subspace is a subset of a vector space that is itself a vector space under the same operations. For example, a line through the origin in $\mathbb{R}^2$ forms a subspace.



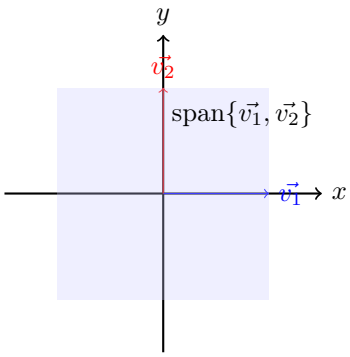
Definition 6. ***V*** *is a vector space over* F, $u_1, u_2, ..., u_n \in V, \lambda_1, \lambda_2, ..., \lambda_n \in F$ *then* $\sum_{i=1}^{n} \lambda_i u_i$ *is called a linear combination. The set of all linear combinations of* $(u_1, ..., u_n)$ *is called the span (linear hull) of* $(u_1, ..., u_n)$.

Notation: $span(u_1, u_2, ..., u_n) := \{\sum_{i=1}^{n} \lambda_i u_i \mid \lambda_i \in F\}$

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

Visual Representation of Linear Combination & Span

Explanation: For example, the span of two vectors $\vec{v_1}$ and $\vec{v_2}$ in $\mathbb{R}^2$ fills a plane.



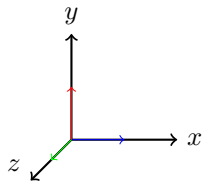
Definition 7. ***A*** *set of vectors* $v_1, ..., v_n$ *is called linearly independent if the following holds:* $\sum_{i=1}^{n} \lambda_i v_i = 0 \Rightarrow \lambda_1 = \lambda_2 = ... = \lambda_n = 0$

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 $\mathbb{R}^d$ is linearly dependent

Visual Representation of Linear Independence

Explanation: In $\mathbb{R}^3$, three vectors pointing along the x, y, and z axes are linearly independent.



Basis and Dimension

Vector spaces are closed under linear combination, and the span of a vector set $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2, \dots\}$ defines a vector subspace. If the entire set of vectors can be obtained by a spanning set, $V = \text{span}(\mathcal{B})$, extending $\mathcal{B}$ by an additional element $\mathcal{C} = \mathcal{B} \cup \{\mathbf{b}\}$ would be redundant since $\text{span}(\mathcal{B}) = \text{span}(\mathcal{C})$. This is recognized by the concept of a basis and also allows for a characterization of the size of a vector space by the cardinality of a basis set.

Definition 8. ***A*** *subset B of a vector space V is called a (Hamel) basis if:*

(P1) $\text{span}(B) = V$ (even for infinite dimensional spaces)

(P2) B is linearly independent

Examples

- The canonical basis of $\mathbb{R}^3$: $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$ (Basis is not unique)
- Another basis of $\mathbb{R}^3$: $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}$ or $\begin{pmatrix}0.5\\0.8\\0.4\end{pmatrix}, \begin{pmatrix}1.8\\0.3\\0.3\end{pmatrix}, \begin{pmatrix}-2.2\\-1.3\\3.5\end{pmatrix}$

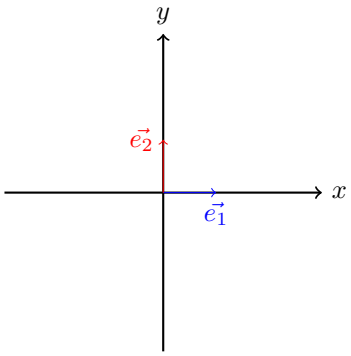
Proposition 9. ***If*** *$U = \{u_1, ..., u_n\}$ spans a vector space V, then the set U can be reduced to a basis of V.*

Informal Proof:

- If U is already linearly independent, done
- If U is dependent: $\exists a \in U$ that is a linear combination of the other vectors in U. We will remove it.
- Keep removing vectors until remaining vectors are linearly independent.

Visual Representation of Basis

Explanation: A basis of a vector space is a set of linearly independent vectors that span the entire space. For $\mathbb{R}^2$, the standard basis consists of $\vec{e_1}$ and $\vec{e_2}$, which point along the x and y axes.



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

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

Proof (sketch): Let $w_1, w_2, ..., w_m$ be a basis of 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 V
- remaining set is linearly independent by construction
- remaining set contains U

Note: For infinite-dimensional spaces, need Zorn's Lemma to prove proposition.

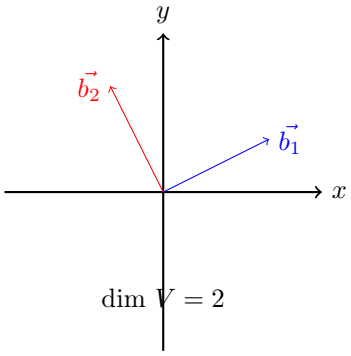
Corollary 12. *Let V be a finite-dimensional vector space, then any two bases of V have the same length.*

Definition 13. *The length of a basis of a finite dimensional vector space is called the dimension of V.*

We have seen what a basis is and what a subspace is. The notion which connects the two is sum and direct sum.

Visual Representation of Finite Dimensions & Dimension of V

Explanation: The dimension of a vector space is the number of vectors in its basis. $\mathbb{R}^2$ has dimension 2 because its basis consists of two vectors.



Definition 14. *Assume that U_1, U_2 are subspaces of V. The sum of the two spaces is defined as,*

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

The sum is called a direct sum, if each element in the sum can be written in exactly one way. Notation: $U_1 \oplus U_2$

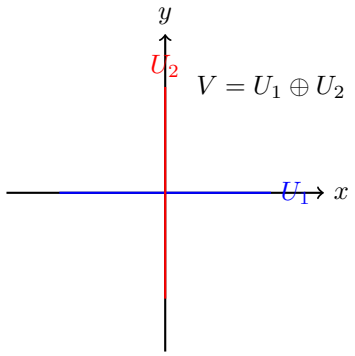
Proposition 15. *Suppose V is a finite-dimensional vector space, and $U \subset V$ is a subspace. Then there exists a subspace $W \subset V$, such that $U \oplus W = V$.*

Proof (sketch): Let the set $\{u_1, u_2, ..., u_n\}$ be a basis of U. Extend it to a basis of V, say the resulting set is $\{u_1, ..., u_n, v_1, v_2, ..., v_m\}$. Define

W = span$\{v_1, v_2, ..., v_m\}$ □

Visual Representation of Direct Sum

Explanation: A vector space V is the direct sum of two subspaces U_1 and U_2 if every vector in V can be uniquely written as the sum of a vector from U_1 and a vector from U_2. In this example, $V = U_1 \oplus U_2$.



In practice the most important procedure to construct direct sums or check when an intersection of two vector subspaces reduces to the zero vector is through an inner product.

Extra concepts that we explored and felt that were relevant to CSE 840

Rings

A ring is a 3-tuple $\mathcal{R} = (R, +, \cdot)$ containing a set R and two operations $+, \cdot$ with properties:

(P1) **Addition Rules:** $(\mathbb{R}, +)$ is a commutative (Abelian) group.

(P2) **Multiplication Rules:**

- Closure: $a \cdot b \in \mathbb{R}$.
- Associativity: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- Identity element: $a \cdot 1 = 1 \cdot a = a$.

(P3) **Distributivity:**

- Left distributive property: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

* Right distributive property: $(a+b)\cdot c = (a\cdot c)+(b\cdot c)$.

As is often the case, a ring is a more complex structure built up from simpler algebraic structures. With respect to addition, a ring has the properties of a commutative group. Only associativity and existence of an identity element is imposed for multiplication. Matrix addition and multiplication has the structure of ring $(\mathbb{R}^{m\times m}, +, \cdot)$.

Proper Subspaces

Definition: A vector subspace must be closed under linear combinations and preserve the same vector addition and scalar multiplication operations as the parent vector space. The simplest example of a vector subspace is the null subspace, $U = \{0\}$. Any subspace must include the zero element since closure requires $u + (-u) = 0$. If $U \subseteq V$, the subspace U is said to be a *proper subspace* of V, denoted by $U < V$.

Setting $n - m$ components equal to zero in the real space $\mathbb{R}_m$ defines a proper subspace whose elements can be placed into a one-to-one correspondence with the vectors within $\mathbb{R}_n$. For example, setting component m of $\mathbf{x} \in \mathbb{R}_m$ equal to zero gives

$$\mathbf{x} = [x_1, x_2, \ldots, x_{m-1}, 0]^T$$

that, while not a member of $\mathbb{R}^n$, is in a one-to-one relation with

$$\mathbf{x}' = [x_1, x_2, \ldots, x_{m-1}]^T \in \mathbb{R}^{m-1}.$$

Dropping the last component of $\mathbf{y} \in \mathbb{R}^m$, where

$$\mathbf{y} = [y_1, y_2, \ldots, y_{m-1}, y_m]^T,$$

gives the vector

$$\mathbf{y}' = [y_1, y_2, \ldots, y_{m-1}]^T \in \mathbb{R}^{m-1},$$

but this is no longer a one-to-one correspondence since, for some given $\mathbf{y}'$, the last component y_m could take any value.

Setting Components to Zero in a Higher-Dimensional Space

When we set $n - m$ components of a vector $\mathbf{x} \in \mathbb{R}^n$ (n-dimensional real space) to zero, we define a subspace of $\mathbb{R}^n$. This subspace corresponds to $\mathbb{R}^m$, the m-dimensional space. Suppose a vector $\mathbf{x} \in \mathbb{R}^m$ is represented as $\mathbf{x} = [x_1, x_2, \ldots, x_{m-1}, 0]^T$, where the last component is set to zero.

This vector is no longer in $\mathbb{R}^n$ but can be mapped one-to-one with $\mathbf{x}' = [x_1, x_2, \ldots, x_{m-1}]^T$, which is in $\mathbb{R}^{m-1}$. If a vector $\mathbf{y} \in \mathbb{R}^m$ is written as $\mathbf{y} = [y_1, y_2, \ldots, y_{m-1}, y_m]^T$, dropping the last component y_m results in $\mathbf{y}' = [y_1, y_2, \ldots, y_{m-1}]^T \in \mathbb{R}^{m-1}$.

However, this process is not a one-to-one mapping because y_m could take any value. Therefore, multiple vectors in $\mathbb{R}^m$ can map to the same $\mathbf{y}'$ in $\mathbb{R}^{m-1}$.

To summarize:

- Setting components to zero can define a proper subspace and create a one-to-one mapping.
- Dropping a component without setting it to zero loses the one-to-one property, as the dropped value can vary freely.

Concept of Orthogonality

Definition: Two vector subspaces U, V of the real vector space $\mathbb{R}^m$ are *orthogonal*, denoted as $U \perp V$, if $\mathbf{u}^T\mathbf{v} = 0$ for any $\mathbf{u} \in U, \mathbf{v} \in V$.

Definition: Two vector subspaces U, V of $U+V$ are *orthogonal complements*, denoted $U = V^\perp$, $V = U^\perp$, if they are orthogonal subspaces, $U \perp V$, and $U \cap V = \{0\}$, i.e., the null vector is the only common element of both subspaces.

The above concept of orthogonality can be extended to other vector subspaces such as spaces of functions It can also be extended to other choices of an inner product in which case the term conjugate vector spaces is sometimes used.

Complement of Direct Sum

If V is a vector space, and U and W are subspaces of V, then:

$$V = U \oplus W$$

implies V is the direct sum of U and W. Here: 1. W is the complement of U (and vice versa) in V. 2. Direct sum conditions: - $U \cap W = \{0\}$: The intersection contains only the zero vector. - Every $v \in V$ can be uniquely written as $v = u + w$, where $u \in U$ and $w \in W$.

Thus, finding a complement W of U in V involves ensuring $V = U \oplus W$. If V is a vector space, and U and W are subspaces of V, then:

$$V = U \oplus W$$

implies V is the direct sum of U and W. Here: 1. W is the complement of U (and vice versa) in V. 2. Direct sum conditions: - $U \cap W = \{0\}$: The intersection contains only the zero vector. - Every $v \in V$ can be uniquely written as $v = u + w$, where $u \in U$ and $w \in W$.

Thus, finding a complement W of U in V involves ensuring $V = U \oplus W$.

References

Sorin Mitran. *Linear Algebra for Data Science*. Retrieved from `http://mitran-lab.amath.unc.edu/courses/MATH347DS/textbook.pdf`.