CPSC 532D — B. FUNCTION SPACES

Danica J. Sutherland

University of British Columbia, Vancouver

Fall 2025

These notes give a quick overview of a few linear algebra concepts that go beyond $\mathbb{R}^d$. For a more thorough introduction, the book of Deisenroth, Faisal, and Ong [DFO20] is focused towards ML, that of Axler [Axl25] is meant as a second course in linear algebra from a more mathematical viewpoint, or that of Roman [Rom07] is a comprehensive graduate-level reference book.

B.1 VECTOR SPACES

You should be already familiar with $\mathbb{R}^d$. You might not be familiar yet with more general notions of vector spaces, which abstract important pieces of the structure of $\mathbb{R}^d$ to let us treat things other than Euclidean vectors in a similar way.

While the definitions can be made more general, we're only going to deal with *real* vector spaces.

Definition B.1. A real *vector space* is a non-empty set V along with the operations of *vector addition*, denoted $v + w \in \mathrm{V}$ for any $v, w \in \mathrm{V}$, and *scalar multiplication*, denoted $av \in \mathrm{V}$ for any $v \in \mathrm{V}$ and $a \in \mathbb{R}$, satisfying the following requirements:

- Vector addition is associative: for all $u, v, w \in \mathrm{V}$, $u + (v + w) = (u + v) + w$.
- Vector addition is commutative: for all $v, w \in \mathrm{V}$, $v + w = w + v$.
- Vector addition has an identity: there is some *zero vector* $0 \in \mathrm{V}$ such that for all $v \in \mathrm{V}$, $v + 0 = v$.
- Vector addition has inverses: for each $v \in \mathrm{V}$, there is some $-v \in \mathrm{V}$ such that $v + (-v) = 0$.
- Compatibility of scalar multiplication: for all $a, b \in \mathbb{R}$ and $v \in \mathrm{V}$, $a(b\mathrm{V}) = (ab)\mathrm{V}$.
- Identity of scalar multiplication: for all $v \in \mathrm{V}$, $(1)v = v$
- Distributive property I: for all $a \in \mathbb{R}$ and $v, w \in \mathrm{V}$, $a(v + w) = av + aw$.
- Distributive property II: for all $a, b \in \mathbb{R}$ and $v \in \mathrm{V}$, $(a + b)v = av + bv$.

The axioms imply many things you'd expect, like that the additive identity is unique, additive inverses are unique and given by $(-1)v$, $0v = 0$ (where the left-hand 0 is in $\mathbb{R}$ and the right-hand is the zero vector in V), etc.

These eight axioms are all things that you probably use all the time without thinking about it for $\mathbb{R}^d$. Given anything with those structures, though, we can do a lot that we do with $\mathbb{R}^d$. For example, we often use *function spaces*. Here $\mathcal{F}$ is some set of functions $\mathcal{X} \to \mathcal{Y}$, and we can treat them as vectors if we define a way to add two

For more, visit https://cs.ubc.ca/~dsuth/532D/25w1/.

functions and to do scalar multiplication, which we will always use as the following: if $f, g \in \mathcal{F}$ and $a \in \mathbb{R}$,

$$f + g = (x \mapsto f(x) + g(x)) \qquad af = (x \mapsto af(x)).$$

These definitions make sense if:

- $\mathcal{Y}$ is itself a vector space, so that adding $f(x) + g(x)$ and multiplying $af(x)$ make sense. This could be just $\mathbb{R}$, or could be more general: $\mathbb{R}^d$ would be common, but it could be itself a function space. . . .

- The set $\mathcal{F}$ needs to be "closed" under both of these operations, so that $f + g$ and af are well-defined. (This is the $v + w \in V$ and $av \in V$ part of the definition.)

So, one reasonable function space could be where $\mathcal{X} = \mathbb{R} = \mathcal{Y}$ and $\mathcal{F}$ is the set of quadratic functions (where here linear or constant functions also count as quadratic). Adding two quadratics gets you another quadratic, as does scaling by a constant. More generally, $\mathcal{F}$ could be the space of polynomials with degree at most k. Much more generally, we can also think about spaces like "all continuous functions" or "all infinitely-differentiable functions."

Definition B.2. A *subspace* U of a real vector space V is a nonempty subset $U \subseteq V$ such that for all $v, w \in U$ and $a \in \mathbb{R}$, $v + w \in U$ and $av \in U$.

The *sum* of two subspaces U, W of V, given by $U + W = \{u + w : u \in U, w \in W\}$, is itself a subspace, and is the smallest subspace containing both U and W. The sum of multiple subspaces is analogous: $\sum_i U_i = \{\sum_i u_i : \forall i,\ u_i \in U_i\}$.

Pairwise independence isn't enough: consider $\operatorname{span}\{(1,0)\}$, $\operatorname{span}\{(0,1)\}$, $\operatorname{span}\{(1,1)\}$ *in* $\mathbb{R}^2$.

A *direct sum*, denoted $\bigoplus_i U_i$ or $U_1 \oplus \cdots \oplus U_m$, is a sum which further satisfies that $U_i \cap (\sum_{j \neq i} U_j) = \{0\}$ for each i. If $V = U \oplus W$, then W is called a *complement* of U.

Note that this implies that any subspace must contain 0 (why?), and that for any V both $\{0\}$ and V are subspaces of V.

Linear functions are a subspace of the space of quadratic functions. More generally, the space of polynomials with degree at most k is a subspace of the space of polynomials with degree at most K for any $k \leq K$, and is also a subspace of the space of all continuous functions.

Something like $\{f \in \mathcal{F} : |f(0)| \leq 1\}$, however, is not a subspace unless $f(0) = 0$ for all $f \in \mathcal{F}$, since a subspace containing f must also contain e.g. $\frac{3}{f(0)} f$.

Any subspace has at least one complement; they are generally not unique, but are isomorphic to one another.

Definition B.3. The *span* of a set of vectors $S \subseteq V$ is the subspace

$$\operatorname{span}(S) = \{a_1 v_1 + \cdots + a_m v_m : m \in \mathbb{N}, a_1, \ldots, a_m \in \mathbb{R}, v_1, \ldots, v_m \in S\}.$$

We define $\operatorname{span}(\{\}) = \{0\}$.

It can be seen that span(S) is the intersection of all subspaces of V containing S.

Definition B.4. A nonempty set of vectors $S \subseteq V$ is *linearly independent* if for any distinct $v_1, \ldots, v_m \in S$,

$$a_1 v_1 + \cdots + a_m v_m = 0 \text{ implies } a_1 = \cdots = a_m = 0.$$

The set {} is defined to be linearly independent. A set which is not linearly independent is *linearly dependent*.

A *basis* for V is a linearly independent set which spans V.

The *dimension* of a vector space V, dim V, is the size of the smallest basis for V; this may be a natural number, or may be infinite.

Every vector space has a basis, and thus has a unique, well-defined dimension. Note that $\dim\{0\} = 0$, while any other vector space has dimension at least one. It can also be seen that if $|S| > \dim V$, then S is linearly dependent.

B.1.1 *Linear Maps*

Definition B.5. Let V and W be real vector spaces. A *linear map* or *linear operator* is a function $T : V \to W$ satisfying that for all $v, v' \in V$ and $a \in \mathbb{R}$,

Some authors use "linear operator" only for $V \to V$ *maps.*

$$T(av + v') = aT(v) + T(v').$$

We also frequently write Tv rather than $T(v)$.

If $\mathcal{F}$ is the function space of infinitely differentiable functions $\mathbb{R} \to \mathbb{R}$, then the derivative operator $Df = (x \mapsto f'(x))$ is a linear operator $\mathcal{F} \to \mathcal{F}$. If $\mathcal{F}_k$ is the space of functions with k derivatives, then $D : \mathcal{F}_k \to \mathcal{F}_{k-1}$ is a linear map.

Definition B.6. A bijective linear map $T : V \to W$ is called an *isomorphism*. If an isomorphism exists, we say that V and W are *isomorphic*.

It turns out that two vector spaces are isomorphic if and only if they have the same dimension. Note that not all infinite-dimensional vector spaces are isomorphic, because they could have different cardinalities, e.g. being countable or uncountable; we won't need to worry about this.

The space of all linear maps from V to W is itself a vector space.

Definition B.7. Let $T : V \to W$ be linear. The *null space* of T is $\operatorname{null}(T) = \{v \in V : Tv = 0\}$, a subspace of V; the *image* of T is $\operatorname{image}(T) = \{Tv : v \in V\}$, a subspace of W. The *rank* of T is $\operatorname{rank}(T) = \dim \operatorname{image}(T)$.

The celebrated "rank-nullity theorem" says that $\operatorname{rank}(T) + \dim \operatorname{null}(T) = \dim V$.

B.2 NORMED SPACES

One thing that our vector spaces don't yet have is a notion of magnitude. In $\mathbb{R}^d$, we usually use the Euclidean norm $\|x\| = \sqrt{\sum_i x_i^2}$, but you've also probably seen at least $\|x\|_1 = \sum_i |x_i|$ and $\|x\|_\infty = \max_i |x_i|$. What properties do these have in common?

Definition B.8. A real *normed vector space* is a real vector space V with a *norm*: a function $V \to \mathbb{R}$, written $\|v\|$, such that:

- Non-negativity: for all $v \in V$, $\|v\| \geq 0$.
- Positive definitenesss: for every $v \in V$, $\|v\| = 0$ if and only if $v = 0$.
- Absolute homogeneity: for every $a \in \mathbb{R}$ and $v \in V$, $\|av\| = |a|\|v\|$.
- Sub-additivity / triangle inequality: for every $v, w \in V$, $\|v + w\| \leq \|v\| + \|w\|$.

The norm of a normed vector space induces the metric $\rho(x, y) = \left\|x - y\right\|$; all the requirements of a metric are easy to see from the properties of a norm.

Definition B.9. Consider a sequence $x_1, x_2, \dots$ in a metric space $\mathcal{X}$.

This sequence has a *limit* x_∞ if for every $\varepsilon > 0$, there exists a positive integer N such that for all $n > N$, $\rho(x_n, x_\infty) < \varepsilon$.

This sequence is called *Cauchy* if, for every $\varepsilon > 0$, there exists a positive integer N such that for all $m, n > N$, $\rho(x_m, x_n) < \varepsilon$.

The metric space $\mathcal{X}$ is called *complete* if all Cauchy sequences in $\mathcal{X}$ have limits in $\mathcal{X}$.

Definition B.10. A real *Banach space* is a real normed vector space whose norm induces a complete metric space.

You can check that, for example, the space of all continuous functions from $\mathcal{X} \to \mathbb{R}$ is a Banach space.

B.3 INNER PRODUCT SPACES

There's one other major structure in $\mathbb{R}^d$ that we don't have yet: dot products.

Definition B.11. A real *inner product space* is a real vector space V together with an inner product, a function $V \times V \to \mathbb{R}$ written $\langle v, w \rangle$ satisfying

- Symmetry: for all $v, w \in V$, $\langle v, w \rangle = \langle w, v \rangle$.
- Linearity: for all $u, v, w \in V$ and $a, b \in R$, $\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$.
- Positive-definiteness: if $v \neq 0$, then $\langle v, v \rangle > 0$.

An inner product space is also a normed vector space with $\|v\| = \sqrt{\langle v, v \rangle}$, and hence a metric space with $\rho(v, w) = \|v - w\| = \sqrt{\langle v - w, v - w \rangle}$.

Definition B.12. A real *Hilbert space* is a real inner product space whose induced metric space is complete.

The usual example of a Hilbert space on functions (...almost ...) is L_2, which has the usual vector space operations and the inner product

$$\langle f, g \rangle_{L_2} = \int f(x) g(x) \, \mathrm{d}x.$$

A slight variant is $L_2(P)$, where P is a probability distribution on $\mathcal{X}$; this has inner product

$$\langle f, g \rangle_{L_2(P)} = \mathop{\mathbb{E}}_{X \sim P} f(X) g(X).$$

The reason this is "almost" a function space is that these functions are only defined almost everywhere: if P is standard normal, and $f' = \begin{cases} f(x) & x \neq 0 \\ 100 & x = 0, \end{cases}$ then $\|f - f'\| = 0$, all inner products are the same, etc. So we usually think of these as being spaces of equivalence classes of functions, rather than of functions proper.

The other major kind of function space is reproducing kernel Hilbert spaces, which we spend a while on later in the course.

REFERENCES

[Axl25] Sheldon Axler. *Linear Algebra Done Right*. 4th edition. 2025.

[DFO20] Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for Machine Learning*. 2020.

[Rom07] Steven Roman. *Advanced Linear Algebra*. 3rd edition. 2007.