

Infinite-dimensional vector spaces

Now that we know what Malliavin Calculus wants to achieve, let us jot down some ideas. If you look for the definition of Malliavin Calculus in Wikipedia, you'll quickly find that there's a Cameron-Martin direction in which the derivative takes place. That doesn't make any sense, so we need to explain what a Cameron-Martin space is. But that requires us to go, before that, to vector spaces with infinite dimensions. And that will require talking about measuring, especially lengths and distances between vectors, and functions as vectors. You'll see I take a lot of inspiration from Slater (2023) and Alessandra Lunardi (2015), great websites that you should visit. This section is quite long and touches many algebra-related topics, so brace yourself.

Hierarchy of spaces

We start from the beginning, with spaces, especially spaces over \mathbb{R} , the field of reals¹. A space is a set V , with elements that we will call vectors, that comes equipped with two operations:

- A “vector addition” operation that returns another vector. That is, $V_1 + V_2 = V_3$
- A “scalar multiplication” operation that returns another vector. That is, $k * V_1 = V_2$

We use the quotes because we don't want to impose any definitions on what those operations are. At most, we will say that they should observe properties like commutation ($V_1 + V_2 = V_2 + V_1$). In practice, though, we won't be so exotic.

Now, we can also equip spaces with other operations that will allow us to measure distances and lengths.

Metric spaces

A metric space is a vector space that's paired with a **metric** or **distance** function $d(V, V) \rightarrow \mathbb{R}$. That is, a function that takes two vectors and gives back a real, positive value. Not every function will do, though. It needs to:

- Be symmetric, that is $d(x, y) = d(y, x)$

¹A word of warning: some of these properties don't apply for the field of complex numbers \mathbb{C} , but we won't deal with them.

- Return a positive value for different vectors, or zero if both vectors are the same
- Have the “triangle inequality”. That is, the distance between two vectors is larger (or equal) if you pass through a third, intermediate vector.

The three most famous distances are the euclidean distance or d_2 , the taxicab distance or d_1 , and the maximum distance or d_∞ . These are their definitions for vectors composed of n real numbers:

$$d_2(\vec{a}, \vec{b}) = \sqrt{\sum_{k=1}^n (a_k - b_k)^2}$$

$$d_1(\vec{a}, \vec{b}) = \sum_{k=1}^n |a_k - b_k|$$

$$d_\infty(\vec{a}, \vec{b}) = \max_k |a_k - b_k|$$

Normed vector spaces

If you pair a vector space with a **norm**, it becomes a normed vector space. A norm is an operation that calculates the length of a single vector and it's denoted as $\|\cdot\| : V \rightarrow \mathbb{R}$. As a function, it takes a single vector and gives back a real, positive value. Again, not every function will do, it needs to:

- Return zero for the zero-vector, or a positive value for non-zero vectors
- Be homogeneous. That is, a vector with components twice as big will have a length twice as large, as in $\|k \cdot x\| = |k| \cdot \|x\|$
- Have the triangle inequality. That is, the norm of a sum is smaller or equal than the sum of the norms

A classical example of norms are the family of p -norms. For a vector of n -length:

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^n (|x_i|)^p}$$

Similar to above, the 1-norm is the taxicab norm $\|x\|_1 = \sum_{i=1}^n |x_i|$, the 2-norm is the euclidean norm and the infinite-norm is the maximum norm. That similarity doesn't end there: you can create a distance out of a norm if you define the distance function to be the norm of the vectors' difference. For example, $d_1(a, b) = |a - b|_1$. This is known as a metric or distance induced by the norm.

The opposite isn't true, though. In general, a distance function can't be turned into a norm, unless:

- The metric is invariant to translations, that is, $d(x, y) = d(x + a, y + a)$
- The metric is homogeneous, that is, $d(kx, ky) = |k| \cdot d(x, y)$

In that case, the metric is induced by the norm $\|x\| = d(x, 0)$.

Inner product spaces

There is yet another step in this ladder. You can equip a vector space with an operation called the **inner product**. It's denoted $\langle V, V \rangle \rightarrow \mathbb{R}$ and it is, like the metric, a function that takes two vectors and gives back a real number. It can be thought of as the length of a vector when you use another vector as a ruler. It has a few requirements, though:

- It must be symmetric, that is, $\langle x, y \rangle = \langle y, x \rangle$
- It must be linear in the first argument, that is, $\langle ix + jy, z \rangle = i\langle x, z \rangle + j\langle y, z \rangle$. It will also apply for the second argument because of symmetry.
- If x is zero, $\langle x, x \rangle = 0$, otherwise it's a strictly positive number.

Every inner product will induce a canonical norm $\|x\| = \sqrt{\langle x, x \rangle}$, which in turn induces a metric. The most common example is the dot product. A n -size vector would be:

$$\langle x, y \rangle = x^T y = \sum_{i=1}^n (x_i \cdot y_i)$$

Infinite-dimensional vector spaces

In the formulas above, we say that there are n components in our vectors. If we want to make sense of Cameron Martin spaces, we need to compare distances and lengths of vectors of potentially infinite dimensions. This is going to be heavily inspired on Slater (2023).

Now, I'm going to assume you know a bit about linear algebra. For example, let's take a vector space with vectors of this form:

$$(a \ b \ c), \ a, b, c \in \mathbb{R}$$

We say that this vector space has dimension 3 because the smallest number of vectors we need to create a linear combination of every vector in that space is 3. Those vectors conform a **basis**, like so:

$$\begin{aligned} (a \ b \ c) &= a(1 \ 0 \ 0) + \\ &\quad b(0 \ 1 \ 0) + \\ &\quad c(0 \ 0 \ 1) \end{aligned}$$

Let's jump to a more interesting case: a vector space that represents polynomials, up to grade $x^n, n \in \mathbb{N}$. That means,

$$a_0 + a_1x + \dots + a_{n-1}x^{n-1} + a_nx^n \rightarrow (a_0 \ a_1 \ \dots \ a_{n-1} \ a_n)$$

We can see that polynomials admit multiplication by a scalar, as well as addition between two polynomials. So, we can truly use these as vectors from a space instead of the full polynomial's formula.

Now, there's nothing preventing us from saying that n is just the *entire* \mathbb{N} , and that the basis is infinite in nature. That is, we need infinitely many vectors to represent all possible polynomials:

$$\begin{aligned} a_0 + a_1x + a_2x^2 + \dots \rightarrow (a_0 \ a_1 \ a_2 \ \dots) &= a_0(1 \ 0 \ 0 \ \dots) + \\ &\quad a_1(0 \ 1 \ 0 \ \dots) + \\ &\quad a_2(0 \ 0 \ 1 \ \dots) + \dots \end{aligned}$$

The idea, while strange, isn't *too* exotic. In fact, a lot of what we already knew for vectors with finite dimensions translate to infinite dimensions. For example, we can apply a linear transformation by multiplying these vectors with an equally-infinite matrix. You may already know that the derivative is a linear operation, so we can represent it with a matrix acting on our vector-as-polynomial:

$$\frac{\partial}{\partial x} (a_0 + a_1x + a_2x^2 + \dots) \rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \dots \\ 0 & 0 & 0 & 0 & \dots \end{pmatrix} * \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ \dots \end{pmatrix} \rightarrow a_1 + 2a_2x + 3a_3x^2 + \dots$$

Infinite sequences

A situation that's a bit more interesting is when evaluating the norm or "length" of an element of these spaces. Let's follow Alessandra Lunardi (2015) and consider the vector space of all infinite sequences of real numbers, (a_i) , which map to \mathbb{R}^∞ . When dealing with finite dimensions, norms will always give you a finite result. This is no longer the case with infinite dimensions. For example, let's take the harmonic sequence:

$$(b_k)_{k \in \mathbb{N}}, b_k = \frac{1}{k} \rightarrow (1 \quad \frac{1}{2} \quad \frac{1}{3} \quad \dots)$$

With the 1-norm,

$$\|b\|_1 = \lim_{n \rightarrow \infty} \sum_{k=1}^n |b_k| = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k} \rightarrow \infty$$

we can't calculate a length value for the sequence, the harmonic series diverges. On the other hand, the 2-norm or euclidean norm, is capable to calculate a value because the series converges to a real number:

$$\|b\|_2 = \lim_{n \rightarrow \infty} \sqrt{\sum_{k=1}^n b_k^2} = \sqrt{\lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k^2}} = \sqrt{\frac{\pi^2}{6}}$$

In fact, mathematicians name ℓ^p the space of sequences where the p -norm converges to a value. Alternatively, if you don't want to deal with norms, you are better off defining a distance function for the space like this:

$$d(x, y) = \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{|x_k - y_k|}{1 + |x_k - y_k|}$$

This will calculate a distance between any two sequences and it will always converge, but this metric isn't translation invariant, and therefore doesn't have an equivalent norm.

Function spaces

Up until now, we have seen countably infinite dimensions, so let's explore an uncountable one now: a space of *functions*. Yeah, functions are now elements of a vector space.

To make any sense of this, we will start with a vector space with an inner product, and we will get the norm and metric for free. Let's remember that for n -size vectors, we did $x^T y$. So, in infinite dimensions, it would be something like:

$$(1 \quad 3 \quad -2 \quad \dots) * \begin{pmatrix} 0 \\ -1 \\ 3 \\ \dots \end{pmatrix} = 1 \cdot 0 + 3 \cdot (-1) + (-2) \cdot 3 + \dots = \lim_{n \rightarrow \infty} \sum_{i=1}^n (x_i \cdot y_i)$$

As we transition from \mathbb{N} to \mathbb{R} , these sums will become integrals. In our vector space of functions defined over some generic domain X :

$$\langle f, g \rangle = \int_X f(x)g(x)dx$$

This definition fulfills our conditions for an inner product, and therefore induces a norm and a distance function:

$$\|f\| = \left(\int_X [f(x)]^2 dx \right)^{\frac{1}{2}}$$

$$d(f, g) = \left(\int_X [f(x) - g(x)]^2 dx \right)^{\frac{1}{2}}$$

As a final note, you can check that there aren't a lot of functions that will give you a value for these as well. Just like ℓ^p restricted the space to sequences that returned a value under the p -norm, we can also restrict the space of functions to those that return a value under the special 2-norm that we induced above. The space is, creatively named, L^p . Alternatively, L^p is named the space of functions that are Lebesgue-integrable.

This space is really very limited when $X = \mathbb{R}$. There aren't many functions where $\int_{-\infty}^{+\infty} [f(x)]^2 dx < \infty$. Only some functions like e^{-x^2} belong there. Polynomials, e^x or logarithms aren't included. That's why people typically define the inner product over a smaller interval, $X = [a, b]$. A norm such as $\int_a^b [f(x)]^2 dx$ admits many more functions and it's much more useful.

There's an additional option that can maintain $X = \mathbb{R}$ and still give you a value: we discard the Lebesgue measure and switch to Gaussian measure. In terms of the Riemann integral, it means that we stop treating all values of X the same and we add a "weighting factor" that will reduce the function values as they go to infinity.

$$\int_{-\infty}^{+\infty} f(x)g(x)dx = \int_{\mathbb{R}} [f \cdot g] d\lambda \rightarrow \int_{\mathbb{R}} [f \cdot g] d\gamma_{\mu, \sigma} = \int_{-\infty}^{+\infty} f(x)g(x)e^{-\frac{(x-\mu)^2}{\sigma}} dx$$

This is the alternative we will choose. As we will see soon, it won't come free of charge.

Linear operators and an alternative derivative definition

Before we continue to the Cameron Martin space, this is probably the best moment to present linear operators. We briefly mentioned them in our polynomials-as-vectors example. We mentioned that we could create a matrix (of infinite size) to represent the derivative. In truth, these matrices can also be thought as functions that transform a vector to another vector.

If this transformation is linear, then the matrix/function is called, unsurprisingly, a linear operator or linear map.

We will use this fancy way of calling matrices/functions to define a derivative in the sense of Fréchet. Fréchet wanted to extend the notion of derivative to work with functions that take m variables as input and outputs an n -sized vector, with our classical derivative being the case for $m = 1, n = 1$. Here's the definition: if you have two normed vector spaces V and W , and a subset $U \subseteq V$. Then $f : U \rightarrow W$ is Fréchet-differentiable at $x \in U$ if there's a linear operator $A : V \rightarrow W$ such that

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(x+h) - f(x) - Ah\|_W}{\|h\|_V} = 0$$

and $A = Df(x)$ is the Fréchet derivative. This may look daunting, but notice that with some minor algebra and alterations, we can obtain a more familiar and straightforward formula:

$$f(x+h) = f(x) + Ah$$

So, in the end, A is how much the function changes with a small h displacement.

This special definition of derivative even allows you to calculate the derivative of the norm $\|\cdot\| : H \rightarrow \mathbb{R}$ around $x \neq 0$:

$$D_x v = \left\langle v, \frac{x}{\|x\|} \right\rangle$$

That is, the derivative of the norm around x , applied on a vector, is how much the length of the vector grows on the direction of x , using a vector of length 1 as a ruler.

Adjoint operators

A final topic around operators is the adjoint operator. Let's take a linear operator $A : U \rightarrow V$, vectors $u \in U, v \in V$, and an inner product for U and V . Then the adjoint of A , called $A^* : V \rightarrow U$, is such that:

$$\langle Au, v \rangle_V = \langle u, A^*v \rangle_U$$

It's important to note that A^* isn't an inverse, although $A^{**} = A$, it's more like a "companion" operator that can be used instead of A if we are in the other vector space.

Orthonormalization

We mentioned above that a basis already has the property of generating every element in the space by linear combination of its (linearly independent) elements. It can also have additional properties.

A basis is “orthogonal” if $\langle b_i, b_j \rangle = 0$ if $i \neq j$. That is, different elements in the basis don’t “overlap” with each other. This means we can express any vector v as a sum that depends only on the basis vectors and v . In particular, if B is an orthogonal basis of V , then any element v of V can be written as:

$$v = \sum_{b_i \in B} a_i b_i = \sum_{b_i \in B} \underbrace{\frac{\langle b_i, v \rangle}{\|b_i\|^2}}_{a_i} b_i$$

With a non-orthogonal basis, you can still write it as a sum, but you need to solve a system of equations to obtain the a_i for the given vector. Compare that to an orthogonal basis, where you can get them directly from an operation that uses the vector and the basis elements themselves.

This can be even better if the basis is “orthonormal”. A basis is orthonormal if it’s orthogonal and $\langle b_i, b_i \rangle = \|b_i\|^2 = 1$. In that case it’s even more straightforward, you really only need a single inner product operation to obtain the scalar that belongs to the basis element:

$$v = \sum_{b_i \in B} a_i b_i = \sum_{b_i \in B} \underbrace{\langle b_i, v \rangle}_{a_i} b_i$$

You may ask why don’t we just use orthonormal basis all the time and save us the hassle. Indeed, we can transform a basis into an orthonormal basis by using the Gram-Schmidt process. I won’t explain the process, it’s only important that we know it can be done.

Alessandra Lunardi, Diego Pallara, Michele Miranda. 2015. “Infinite Dimensional Analysis.” 2015. <http://www.dm.unife.it/it/ricerca-dmi/seminari/sem19/lectures/lecture-notes/view>.

Slater, Max. 2023. “Functions Are Vectors.” 2023. <https://thenumb.at/Functions-are-Vectors>.