The Malliavin Derivative

Liam Llamazares

06/30/2022

1 Three line summary

- The Malliavin derivative is an operator defined by manipulating the chaos expansion of a square-integrable random variable.
- The Malliavin derivative transforms square-integrable variables into square integrable processes.
- The Malliavin derivative shares some properties with the classic derivative such as the product rule and chain rule, but the fundamental theorem of calculus only holds in some special cases.

2 Why should I care?

The Malliavin derivative is (somewhat unsurprisingly) a fundamental object of Malliavin calculus and has many applications in finance, numerical methods, and optimal control.

3 Notation

The same as in previous posts. Furthermore we will shorten the notation $L^2(\Omega, \mathcal{F}_T)$ and $L^2(I \times \Omega, \mathcal{B}(I) \otimes \mathcal{F}_T)$ to $L^2(\Omega)$ and $L^2(I \times \Omega)$ respectively.

4 Introduction

The Malliavin derivative was originally introduced as an operator associated with the Fréchet differential of random variables $X : C(I) \to \mathbb{R}$. The aforementioned construction provides some motivation behind the Malliavin derivative and will be developed in the next post. However, a more general construction can be obtained via the chaos expansion.

Definition 1. *Given* $X = \sum_{n=0}^{\infty} I_n(f_n) \in L^2(\Omega)$ *we say that* X *is* Malliavin differentiable *if*

$$\|X\|_{\mathbb{D}_{1,2}} := \sum_{n=1}^{\infty} nn!\|f_n\|_{L^2([0,T]^n)} < \infty,$$

and denote the space of Malliavin differentiable functions by

$$\mathbb{D}_{1,2} := \{X \in L^2(\Omega) : \|X\|_{\mathbb{D}_{1,2}} < \infty\}.$$

Furthermore, we define the Malliavin derivative *of* X *as*

$$D_t X := \sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot, t)),$$

The first questions to be asked is: "what kind of object is the Malliavin derivative of a random variable? What is the link between $\mathbb{D}_{1,2}$ and D_t? Is $\|\cdot\|_{\mathbb{D}_{1,2}}$ even a norm?" We answer this in the next proposition and corollary.

Proposition 1. *The Malliavin derivative is well defined on* $\mathbb{D}_{1,2}$ *and establishes a linear isometry*

$$\begin{aligned} D : (\mathbb{D}_{1,2}, \|\cdot\|_{\mathbb{D}_{1,2}}) &\longrightarrow L^2(I \times \Omega, \|\cdot\|_{L^2(I\times\Omega)}) \\ X &\longmapsto DX. \end{aligned}$$

Proof. The proof of the first part of the proposition is a straightforward application of Itô's n-th isometry and the monotone convergence theorem as we have that

$$\begin{aligned} \|DX\|^2_{L^2(I\times\Omega)} &= \int_I \left\| \sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot, t)) \right\|^2_{L^2(\Omega)} dt \\ &= \sum_{n=1}^{\infty} n^2(n-1)! \int_I \|f_n(\cdot, t)\| dt = \sum_{n=1}^{\infty} nn!\|f_n\|_{L^2([0,T]^n)} = \|X\|_{\mathbb{D}_{1,2}}. \end{aligned}$$

Finally, the linearity of D follows from the linearity of the iterated Itô integrals (which itself is a consequence of the linearity of the Itô integral). $\square$

In summary, the Malliavin derivative turns a square-integrable random variable into a possibly non-adapted, stochastic process. You may recall from our previous posts that we had an operator that went in the opposite direction. The Skorohod integral δ. In fact, the Malliavin derivative and the Skorohod integral are adjoint operators in a sense that will be made precise in the next post. For now, we show that, as occurs with the ordinary derivative, a random variable has Malliavin derivative 0 if and only if it is constant.

Corollary 1. $(\mathbb{D}_{1,2}, \|\cdot\|_{\mathbb{D}_{1,2}})$ *is a seminormed space. Furthermore*

$$\|X\|_{\mathbb{D}_{1,2}} = 0 \iff D_t X = 0 \iff X \in \mathbb{R}.$$

Proof. The triangle inequality and the absolute homogeneity are direct consequences of the isometry of the previous proposition. This shows that $\|\cdot\|_{\mathbb{D}_{1,2}}$ is a seminorm. The second part follows from the fact that

$$\|X\|_{\mathbb{D}_{1,2}} = \sum_{n=1}^{\infty} n n! \|f_n\|_{L^2([0,T]^n)}.$$

So $\|X\|_{\mathbb{D}_{1,2}} = 0$ if and only if $f_n = 0$ for all $n \geq 1$, which in turn is equivalent to $X = I_0(f_0) := f_0$. Where we recall that by convention $L^2_S(I^0) := \mathbb{R}$ and I_0 was defined as the identity on $\mathbb{R}$. This concludes the proof. $\square$

Before moving on we show a motivating example. Let us consider some deterministic function $f \in L^2(I)$ and set

$$X := \delta(f) = \int_0^T f(s) dW(s).$$

Then, by construction, we have $X = I_1(f)$ so

$$D_t X = I_0(f(t)) = f(t).$$

This is a nice result and it might suggest something akin to the fundamental theorem of calculus such as $D_t(\delta Y) = Y$ for any Skorohod integrable process Y. However, this will not hold in general and as, will be seen in the next post, occurs if and only if Y is a deterministic function in $L^2(\Omega)$.

This said, we now show that $\mathbb{D}_{1,2}$ is closed in the sense that: given a convergent sequence $X_m \to X$. If the derivatives $D_t X_m$ converge then also $D_t X_m \to D_t X$.

Proposition 2. *Let $X_m \in \mathbb{D}_{1,2}$ such that X_m is a Cauchy sequence in both $L^2(\Omega)$ and in $\mathbb{D}_{1,2}$. Then, there exists $X \in \mathbb{D}_{1,2}$ such that*

$$\lim_{m\to\infty} \|X_m - X\|_{L^2(\Omega)} = \lim_{n\to\infty} \|D_t X_m - D_t X\|_{L^2(I\times\Omega)}.$$

Proof. First of all, we note that since $L^2(\Omega)$ is complete X_m must converge to some $X \in L^2(\Omega)$. Let us write the respective chaos expansions as

$$X = \sum_{n=0}^{\infty} I_n(f_n); \quad X_m = \sum_{n=0}^{\infty} I_n(f_n^{(m)}).$$

By Itô's isometry and the convergence $X_m \to X \in L^2(\Omega)$ we deduce that also $f_n^{(m)} \to f_n \in L^2(I^n)$ for each n. By now applying Fatou's lemma and the fact that X_m is by hypothesis Cauchy in $\mathbb{D}_{1,2}$ we obtain that

$$\lim_{m\to\infty} \|X - X_m\|_{\mathbb{D}_{1,2}} = \lim_{m\to\infty} \sum_{n=1}^{\infty} n!n\|f_n(x) - f_n^{(m)}\|_{L^2(I^n)}$$

$$\leq \lim_{m\to\infty} \liminf_{k\to\infty} \sum_{n=1}^{\infty} n!n\|f_n^{(k)}(x) - f_n^{(m)}\|_{L^2(I^n)} = \lim_{m\to\infty} \liminf_{k\to\infty} \|X_m - X_k\|_{\mathbb{D}_{1,2}} = 0.$$

As desired. $\square$

We now conclude this post by stating two properties of the Malliavin derivative that are analogous to those verified by the derivative of ordinary functions. Firstly, an analog to the chain rule for the ordinary derivative. The proof can be found on page 29 of Nunno and Øksendal's book [1] but is rather technical and relies on Hermite polynomials which were not discussed previously, so we omit it.

Definition 2. *We write $\mathbb{D}_{1,2}^0 \subset L^2(\Omega)$ for the space of square integrable random variables $X = \sum_{n=0}^{\infty} I_n(f_n)$ such that $f_n = 0$ for all but finitely many n.*

Proposition 3 (Product rule)**.** *Given $X_1, X_2 \in \mathbb{D}^0_{1,2}$ it holds that*

$$D_t(X_1X_2) = X_1D_tX_2 + X_2D_tX_1.$$

Finally, though we shall not use it, we mention that if $\Omega = \mathcal{S}^*(\mathbb{R})$ is the dual of the Schwartz space and we construct a probability measure $\mathbb{P}$ called the *white noise probability measure* then the following version of the chain rule also holds (see [1] page 89).

Proposition 4 (Chain rule)**.** *Consider $\varphi \in C_1(\mathbb{R}^d)$ with $\nabla\varphi \in L^\infty(\mathbb{R}^d \to \mathbb{R}^d)$ and $X = (X_1, \ldots, X_d)$ such that $X_i \in \mathbb{D}_{1,2}$ for each $i = 1, \ldots, d$. Then it holds that $\varphi(X) \in \mathbb{D}_{1,2}$ with*

$$D_t\varphi(X) = \sum_{i=0}^{d} \frac{\partial\varphi}{\partial x_i}(X)D_t(X_i).$$

5 Extending past p=2

The Malliavin derivative lets us define a derivative on a subset of $L^2(\Omega)$. However, it may also be useful to have a concept of derivative on random variables in $L^p(\Omega)$. We now explain how to do this via an alternative construction of the Malliavin derivative. First of all, consider the set of *cylindrical variables*

$$\mathbb{W} := \left\{\varphi\left(\int_I h(t)dW(t)\right) : \varphi \in C_b^\infty(\mathbb{R}^n), h \in L^2(I \to \mathbb{R}^n), n \in \mathbb{N}\right\}$$

That is, $\mathbb{W}$ is the set of all smooth functions with bounded derivatives of Wiener integrals of deterministic functions. Let us use the abbreviation $W(h) := \int_I h(t)dW(t)$. Then, the results in the previous section show that the Malliavin differential of a cylindrical variable is

$$D_t\varphi(W(h)) = \nabla\varphi(W(h)) \cdot h(t) = \sum_{i=1}^{n} \frac{\partial\varphi}{\partial x_i}(W(h))h_i(t).$$

One can also start directly with the above equation as the defintion of Malliavin differential. In this case it is not clear that D_t is well defined (that is, independent of the representation of $X = \varphi(W(h))$). However it is, see [2] page 10. Analogously to how one defines the norm on Sobolev spaces, we now take any $1 \leq p < \infty$ and define a norm on $\mathbb{W}$ by

$$\|X\|_{\mathbb{D}^{1,p}} := \|X\|_{L^p(\Omega)} + \|DX\|_{L^p(\Omega \to L^2(I))}, \quad X \in \mathbb{W}.$$

Then, D is a continuous linear operator on $(\mathbb{W}, \|\cdot\|_{\mathbb{D}^{1,p}})$ to $L^2(\Omega \to L^2(I))$. As a result, D may be extended to the closure of $(\mathbb{W}, \|\cdot\|_{\mathbb{D}^{1,p}})$. We denote this closure by $\mathbb{D}^{1,p}$ and by abuse of notation also write D for the continuous extension of D to $\mathbb{D}^{1,p}$. Note that by definition of the norm $\|\cdot\|_{\mathbb{D}^{1,p}}$, necessarily $\mathbb{D}^{1,p}$ is a subset of $L^p(\Omega)$. In this way, we have been able to extend the Malliavin differential to $\mathbb{D}^{1,p} \subset L^p(\Omega)$. Explicitly, we have that

$$DX := \lim_{n\to\infty} DX_n \in L^p(\Omega \to L^2(I)).$$

Where $X_n \in \mathbb{W}$ is a sequence converging to X in $\mathbb{D}^{1,p}$. Furthermore, we note that by the previous discussion D coincides with our previous definition of the Malliavin differential when $p = 2$. For the case $p = \infty$ we define

$$D^{1,\infty} := \bigcap_{p=1}^{\infty} \mathbb{D}^{1,p}.$$

We now conclude with an extension of the chain rule which can be used even when φ does not have bounded derivative.

Proposition 5 (Chain rule for $\mathbb{D}^{1,p}$)**.** *Let $X \in \mathbb{D}^{1,p}$ and consider $\varphi \in C^1(\mathbb{R}^n)$ such that $\|\nabla\varphi(x)\| \leq C(1+\|x\|^\alpha)$ for some $0 \leq \alpha \leq p-1$. Then $\varphi(X) \in \mathbb{D}^{1,q}$, where $q = p/(\alpha+1)$. Furthermore,*

$$D\varphi(X) = \nabla\varphi(X) \cdot DX.$$

Proof. By the mean value inequality we have that

$$|\varphi(x)| \leq C'(1 + \|x\|^{\alpha+1}) = C'(1 + \|x\|^{\frac{p}{q}}).$$

As a result we have that

$$\varphi(X) \in L^q(\Omega). \tag{1}$$

Furthermore, by Hölder's inequality applied to $r = (\alpha+1)/\alpha, s = \alpha + 1$ we have that

$$\nabla\varphi(X) \cdot DX \in L^q(\Omega \to L^2(I)). \tag{2}$$

We now take a sequence of cylindrical random variables X_n converging to X in $\mathbb{D}^{1,p}$ and an approximation to the identity δ_n. Let us set $\varphi_n := \varphi \cdot \delta_n$

$$\begin{aligned} D\varphi(X) = \lim_{n\to\infty} D[\varphi_n(X_n)] &= \lim_{n\to\infty} (\nabla\varphi_n)(X_n) \cdot DX_n \\ &= (\nabla\varphi)(X) \cdot DX \in L^q(\Omega \to L^2(I)). \end{aligned}$$

Where the final equality is due to the same method that gave inclusions (1)-(2) and the way X_n, φ_n converge to X, φ respectively. $\square$

Essentially the previous proposition says that, if X is differentiable and the derivative of φ doesn't grow to fast (depending on the integrability of DX), then $\varphi(X)$ is also differentiable and we can apply the chain rule. The integrability of $D\varphi(X)$ depending on the integrability of DX and the growth of $\nabla\varphi$.

5.1 Example application

For example, in the case $p = 2$ we could take $\varphi(x) = x^2$ to deduce that

$$D(X^2) = 2XDX, \quad \forall X \in \mathbb{D}^{1,2}.$$

If $X = 1_A$ is an indicator function for some $A \in \mathcal{F}$ we obtain that

$$D(1_A) = D(1_A^2) = 21_A D(1_A).$$

From here we deduce that $D1_A = 0$. As we have seen, this occurs if and only if 1_A is constant, so necessarily $1_A = 0$ or $1_A = 1$. Identifying sets with functions, we have just proved the following

Corollary 2. *Given $A \in \mathcal{F}$ we have that $1_A \in \mathbb{D}^{1,2}$ if and only if (almost everywhere) $A = \emptyset$ or $A = \Omega$.*

5.2 Multiple derivatives

Finally we comment on how it is possible to iterate the Malliavin derivative. Given a cylindrical process $X = \varphi(W(h)) \in \mathbb{W}$ we should have that, with Einstein notation

$$D_{t_1}D_{t_2}X = D_{t_1}(\partial_i\varphi(X)h_i(t_2)) = D_{t_1}(\partial_i\varphi(X))h_i(t_2) = (\partial_i\partial_j\varphi)(X)h_i(t_1)h_j(t_2).$$

As a result, we define given $X \in \mathbb{W}$

$$D^k_{t_1,\ldots,t_k}X := \sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial\varphi}{\partial x_{i_1}\ldots\partial x_{i_k}}(X)h_{i_1}(t_1)\ldots h_{i_k}(t_k).$$

In the same fashion as before, we can now define the k-th differential norm as

$$\|X\|_{\mathbb{D}^{k,p}} := \sum_{i=0}^{k} \left\|D^j X\right\|_{L^p(\Omega\to L^2(I^k))}.$$

Where we use the convention $D^0 X := X, L^2(I^0) := \mathbb{R}$. Then we simply define $\mathbb{D}^{p,k}$ to be the completion of $\mathbb{W}$ with this norm. Finally, we define

$$\mathbb{D}^{\infty,p} := \bigcap_{k=1}^{\infty} \mathbb{D}^{k,p}; \quad \mathbb{D}^{\infty} := \mathbb{D}^{\infty,\infty}.$$

References

[1] G. D. Nunno, B. Øksendal, F. Proske, Malliavin calculus for Lévy processes with applications to finance, Springer, 2008.
URL https://link.springer.com/book/10.1007/978-3-540-78572-9

[2] M. Hairer, Introduction to malliavin calculus (2021).
URL https://www.hairer.org/notes/Malliavin.pdf