Chapter 2

Calculus in Gauss Space

ı:Calc_on_Gauss_Space⟩

In this section we develop the basics of calculus on the finite-dimensional Gauss space. The differences between this calculus and the "regular" calculus that we first learn (which we call calculus on Lebesgue space) are not that stark. At the end of the day we still compute integrals and derivatives in the same way, but there are some modifications that must be taken into account. The most important of these is the *integration-by-parts* formula, which must be modified to properly accomodate for the Gaussian background measure. On a computational level this modification is elementary. But we shall see that it has far-reaching consequences.

1 The Gradient Operator

The *n-dimensional Lebesgue space* is the measurable space $(\mathbb{E}^n, \mathcal{B}(\mathbb{E}^n))$—where $\mathbb{E} = [0\,,1)$ or $\mathbb{E} = \mathbb{R}$—endowed with the Lebesgue measure, and the "calculus of functions" on Lebesgue space is just "real and harmonic analysis."

The n-dimensional *Gauss space* is the same measure space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ as in the previous paragraph, but is now endowed with the Gauss measure P_n in place of the Lebesgue measure. Since the Gauss space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\,, \mathrm{P}_n)$ is a probability space, we can—and frequently will—think of a measurable function $f : \mathbb{R}^n \to \mathbb{R}$ as a random variable. Therefore,

$$\begin{aligned} \mathrm{P}\{f \in A\} &= \mathrm{P}_n\{f \in A\} = \mathrm{P}_n\{x \in \mathbb{R}^n : f(x) \in A\}, \\ \mathrm{E}(f) &= \mathrm{E}_n(f) = \int f\,\mathrm{d}\mathrm{P}_n = \int f\,\mathrm{d}\mathrm{P}, \\ \mathrm{Cov}(f\,,g) &= \langle f\,,g\rangle_{L^2(\mathrm{P})} = \int fg\,\mathrm{d}\mathrm{P}, \end{aligned}$$

etc. Note, also, that $f = f(Z)$ for all random variables f, where Z is the standard normal random vector $Z(x) := x$ for all $x \in \mathbb{R}^n$, as before. In particular,

$$\mathrm{E}(f) = \mathrm{E}_n(f) = \mathrm{E}[f(Z)],$$

$$\mathrm{Var}(f) = \mathrm{Var}[f(Z)], \quad \mathrm{Cov}(f\,,g) = \mathrm{Cov}[f(Z)\,,g(Z)], \ldots$$

and so on, notation being typically obvious from context.

Let $\partial_j := \partial/\partial x_j$ for all $1 \leqslant j \leqslant n$. From now on we will use the following.

Definition 1.1. Let $C_0^k(\mathrm{P}_n)$ denote the collection of all infinitely-differentiable functions $f:\mathbb{R}^n\to\mathbb{R}$ such that f and all of its mixed derivatives of order $\leqslant k$ grow more slowly than $[\gamma_n(x)]^{-\varepsilon}$ for every $\varepsilon>0$. We also define

$$C_0^\infty(\mathrm{P}_n) := \bigcap_{k=1}^\infty C_0^k(\mathrm{P}_n).$$

It is not hard to see that $f\in C_0^k(\mathrm{P}_n)$ if and only if

$$\lim_{\|x\|\to\infty} \mathrm{e}^{-\varepsilon\|x\|^2}|f(x)| = \lim_{\|x\|\to\infty} \mathrm{e}^{-\varepsilon\|x\|^2}\,|(\partial_{i_1}\cdots\partial_{i_m}f)(x)| = 0,$$

for all $1\leqslant i_1,\ldots,i_m\leqslant n$ and $1\leqslant m\leqslant k$ (see Problem 4).

We will frequently use the following result without explicit mention.

⟨lem:Ck_moments⟩ **Lemma 1.2.** *If $f\in C_0^k(\mathrm{P}_n)$, then*

$$\mathrm{E}\left(|f|^p\right)<\infty \quad \textit{and} \quad \mathrm{E}\left(|\partial_{i_1}\cdots\partial_{i_m}f|^p\right)<\infty,$$

for all $1\leqslant p<\infty$, $1\leqslant i_1,\ldots,i_m\leqslant n$, and $1\leqslant m\leqslant k$.

The proof is relegated to Problem 1.

For every $f\in C_0^1(\mathrm{P}_n)$, define

$$\begin{aligned}\|f\|_{1,2}^2 &:= \int |f(x)|^2\,\mathrm{P}_n(\mathrm{d}x) + \int \|(\nabla f)(x)\|^2\,\mathrm{P}_n(\mathrm{d}x)\\ &= \mathrm{E}\left(|f|^2\right) + \mathrm{E}\left(\|\nabla f\|^2\right),\end{aligned}$$

where $\nabla := (\partial_1,\ldots,\partial_n)$ denotes the gradient operator. Notice that $\|\cdot\|_{1,2}$ is a *bona fide* Hilbertian norm on $C_0^1(\mathrm{P}_n)$ with Hilbertian inner product

$$\begin{aligned}\langle f\,,g\rangle_{1,2} &:= \int fg\,\mathrm{d}\mathrm{P}_n + \int (\nabla f)\cdot(\nabla g)\,\mathrm{d}\mathrm{P}_n\\ &= \mathrm{E}[fg] + \mathrm{E}[\nabla f\cdot\nabla g].\end{aligned}$$

We will soon see that $C_0^1(\mathrm{P}_n)$ is not a Hilbert space with the preceding norm and inner product because it is not complete; that is, there are Cauchy sequences in $C_0^1(\mathrm{P}_n)$ that fail to in $C_0^1(\mathrm{P}_n)$. Thus, we are led to the following.

Definition 1.3. The *Gaussian Sobolev space* $\mathbb{D}^{1,2}(\mathrm{P}_n)$ is the completion of $C_0^1(\mathrm{P}_n)$ in the norm $\|\cdot\|_{1,2}$.

In order to understand what the elements of $\mathbb{D}^{1,2}(\mathrm{P}_n)$ look like, let us consider a function $f\in\mathbb{D}^{1,2}(\mathrm{P}_n)$. By definition, we can find a sequence $f_1,f_2,\ldots\in C_0^1(\mathrm{P}_n)$ such that $\|f_\ell - f\|_{1,2}\to 0$ as $\ell\to\infty$. Since $L^2(\mathrm{P}_n)$ is complete, we can deduce also that

$$D_jf := \lim_{\ell\to\infty}\partial_j f_\ell \qquad \text{exists in } L^2(\mathrm{P}_n) \text{ for every } 1\leqslant j\leqslant n.$$

It follows, by virtue of construction, that

$$Df = \nabla f \qquad \text{for all } f\in C_0^1(\mathrm{P}_n).$$

Therefore, D is an extension of the gradient operator from $C_0^1(\mathrm{P}_n)$ to $\mathbb{D}^{1,2}(\mathrm{P}_n)$. From now on, we will almost always write Df in favor of ∇f when $f\in C_0^1(\mathrm{P}_n)$. This is because Df can make sense even when f is not in $C_0^1(\mathrm{P}_n)$, as we will see in the next few examples.

In general, we can think of elements of $\mathbb{D}^{1,2}(\mathrm{P}_n)$ as functions in $L^2(\mathrm{P}_n)$ that have one weak derivative in $L^2(\mathrm{P}_n)$. We may refer to the linear operator D as the *Malliavin derivative*, and the random variable Df as the [generalized] *gradient of* f. We will formalize this notation further at the end of this section. For now, let us note instead that the standard Sobolev space $W^{1,2}(\mathbb{R}^n)$ is obtained in exactly the same way as $\mathbb{D}^{1,2}(\mathrm{P}_n)$ was, but the Lebesgue measure is used in place of P_n everywhere. Since $\gamma_n(x) = \mathrm{dP}_n(x)/\mathrm{d}x < 1$,[1] it follows that the Hilbert space $\mathbb{D}^{1,2}(\mathrm{P}_n)$ is richer than the Hilbert space $W^{1,2}(\mathbb{R}^n)$, whence the Malliavin derivative is an extension of Sobolev's [generalized] gradient. The extension is strict; see Problem 6.

It is a natural time to produce examples to show that the space $\mathbb{D}^{1,2}(\mathrm{P}_n)$ is strictly larger than the space $C_0^1(\mathrm{P}_n)$ endowed with the norm $\|\cdot\|_{1,2}$.

⟨ex:Smoothing:1⟩ *Example* 1.4 ($n = 1$). Consider the case $n = 1$ and let f denote the "tent function," $f(x) := (1 - |x|)_+$ on $\mathbb{R}$. We claim that $f \in \mathbb{D}^{1,2}(\mathrm{P}_1) \setminus C_0^1(\mathrm{P}_1)$. Moreover, we claim

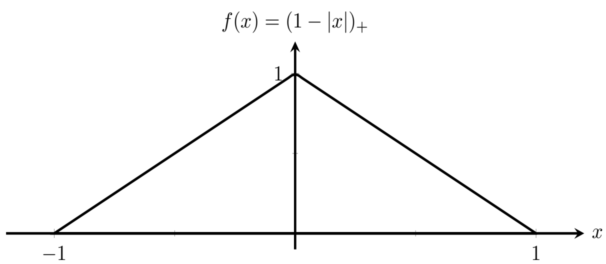

Figure 2.1. A tent function.

?⟨fig:tent⟩?

the P_1-a.s. identity,[2]

$$(Df)(x) = -\mathrm{sign}(x)\mathbb{1}_{[-1,1]}(x).$$

In a sense, this formula is obvious. We propose to derive it rigorously, thus emphasizing the fact that Df should be regarded as an element of $L^2(\mathrm{P}_n)$.

Let $\psi_1 \in C^\infty(\mathbb{R})$ be a symmetric probability density function on $\mathbb{R}$ such that $\psi_1 \equiv$ a positive constant on $[-1\,,1]$, and $\psi_1 \equiv 0$ off $[-2\,,2]$. For every real number $r > 0$, define $\psi_r(x) := r\psi_1(rx)$ and $f_r(x) := (f * \psi_r)(x)$. Then $\sup_x |f_N(x) - f(x)| \to 0$ as $N \to \infty$ because f is uniformly continuous. In particular, $\|f_N - f\|_{L^2(\mathrm{P}_n)} \to 0$ as $N \to \infty$. To complete the proof it remains to verify that

$$\lim_{N\to\infty} \int |f_N'(x) + \mathrm{sign}(x)\mathbb{1}_{[-1,1]}(x)|^2 \, \mathrm{P}_n(\mathrm{d}x) = 0. \tag{2.1}$$ goal:n=1

[1] In other words, $\mathrm{E}(|f|^2) < \int_{\mathbb{R}^n} |f(x)|^2 \, \mathrm{d}x$ for all $f \in L^2(\mathbb{R}^n)$ that are strictly positive on a set of positive Lebesgue measure.

[2] It might help to recall that Df is defined as an element of the Hilbert space $L^2(\mathrm{P}_1)$ in this case. Therefore, it does not make sense to try to compute $(Df)(x)$ for all $x \in \mathbb{R}$. This issue arises when one constructs any random variable on any probability space, of course. Also, note that P_1-a.s. equality is the same thing as Lebesgue-a.e. equality, since the two measures are mutually absolutely continuous.

By the dominated convergence theorem and integration by parts,

$$\begin{aligned} f_N'(x) &= \int_{-\infty}^{\infty} f(y)\psi_N'(x-y)\,\mathrm{d}y \\ &= -\int_0^1 \psi_N(x-y)\,\mathrm{d}y + \int_{-1}^0 \psi_N(x-y)\,\mathrm{d}y \\ &:= \ -A_N(x) + B_N(x). \end{aligned}$$

We now prove that $A_N \to \mathbb{1}_{[0,1]}$ as $N \to \infty$ in $L^2(\mathrm{P}_1)$; a small adaptation of this argument will also prove that $B_N \to \mathbb{1}_{[-1,0]}$ in $L^2(\mathrm{P}_1)$, from which (2.1) ensues.

By a change of variables, $A_N(x) = \int_{N(x-1)}^{Nx} \psi_1(y)\,\mathrm{d}y$. Because ψ_1 is a probability density function, it follows that $A_N(x) \to \mathbb{1}_{[0,1]}(x)$ as $N \to \infty$ for P_1-almost all x. Similarly $B_n(x) \to \mathbb{1}_{[-1,0]}(x)$ for P_1-almost all x, and therefore $f_N'(x) = -A_N(x) + B_N(x) \to -\mathrm{sign}(x)\mathbb{1}_{[-1,1]}(x)$ for P_1-almost all x. Since $f_N'(x) - \mathrm{sign}(x)\mathbb{1}_{[-1,1]}(x)$ is bounded uniformly by 2, the dominated convergence theorem implies that the convergence also takes place in $L^2(\mathrm{P}_1)$. This concludes our example.

⟨ex:Smoothing:2⟩ *Example* 1.5 ($n \geqslant 2$). Let us consider the case that $n \geqslant 2$. In order to produce a function $F \in \mathbb{D}^{1,2}(\mathrm{P}_n) \setminus C_0^1(\mathrm{P}_n)$ we use the construction of the previous example and set

$$F(x) := \prod_{j=1}^n f(x_j) \text{ and } \Psi_N(x) := \prod_{j=1}^n \psi_N(x_j) \quad \text{for all } x \in \mathbb{R}^n \text{ and } N \geqslant 1.$$

Then the calculations of Example 1.4 also imply that $F_N := F * \Psi_N \to F$ as $N \to \infty$ in the norm $\|\cdots\|_{1,2}$ of $\mathbb{D}^{1,2}(\mathrm{P}_n)$, $F_N \in C_0^1(\mathrm{P}_n)$, and $F \notin C_0^1(\mathrm{P}_n)$. Thus, it follows that $F \in \mathbb{D}^{1,2}(\mathrm{P}_n) \setminus C_0^1(\mathrm{P}_n)$. Furthermore,

$$(D_jF)(x) = -\mathrm{sign}(x_j)\mathbb{1}_{[-1,1]}(x_j) \times \prod_{\substack{1 \leqslant \ell \leqslant n \\ \ell \neq j}} f(x_\ell),$$

for every $1 \leqslant j \leqslant n$ and P_n-almost every $x \in \mathbb{R}^n$.

⟨ex:Lipschitz:D12⟩ *Example* 1.6. The previous two examples are particular cases of a more general family of examples. Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is *Lipschitz continuous* if there exists a finite constant K such that $|f(x) - f(y)| \leqslant K\|x - y\|$ for all $x, y \in \mathbb{R}^n$. The smallest such constant K is called the *Lipschitz constant of* f and is denoted by $\mathrm{Lip}(f)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function. According to Rademacher's theorem XXX, f is almost everywhere [equivalently, P_n-a.s.] differentiable and $\|(\nabla f)(x)\| \leqslant \mathrm{Lip}(f)$ a.s. Also note that

$$|f(x)| \leqslant |f(0)| + \mathrm{Lip}(f)\|x\| \qquad \text{for all } x \in \mathbb{R}^n.$$

In particular, $\mathrm{E}(|f|^k) < \infty$ for all $k \geqslant 1$. A density argument, similar to the one that appeared in the preceding examples, shows that $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ and

$$\|(Df)(x)\| \leqslant \mathrm{Lip}(f) \qquad \mathrm{P}_1\text{-almost all } x.$$

We will appeal to this fact several times in this book.

The generalized gradient D follows more or less the same general set of rules as does the more usual gradient operator ∇. And it frequently behaves as one expect it should even when it is understood as the Gaussian extension of ∇; see Examples 1.4 and 1.5, for instance. The following ought to reinforce this point of view.

⟨lem:ChainRule⟩ **Lemma 1.7** (Chain Rule). *For all* $\psi \in \mathbb{D}^{1,2}(\mathrm{P}_1)$ *and* $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$,

$$D(\psi \circ f) = [(D\psi) \circ f]\, D(f) \qquad a.s.$$

Proof. If f and ψ are smooth functions, then the chain rule of calculus ensures that $[\partial_j(\psi \circ f)](x) = \psi'(f(x))(\partial_j f)(x)$ for all $x \in \mathbb{R}^n$ and $1 \leqslant j \leqslant n$. That is,

$$D(g \circ f) = \nabla(\psi \circ f) = (\psi' \circ f)(\nabla f) = (D\psi)(f)D(f),$$

where $D\psi$ refers to the one-dimensional Malliavin derivative of ψ and $D(f) := Df$ refers to the n-dimensional Malliavin derivative of f. The general case follows from the smooth case and a density argument. □

Here is a final example that is worthy of mention.

⟨ex:DM⟩ *Example* 1.8. Let $M := \max_{1 \leqslant j \leqslant n} Z_j$ and note that

$$M(x) = \max_{1 \leqslant j \leqslant n} x_j = \sum_{j=1}^{n} x_j \mathbb{1}_{Q(j)}(x) \qquad \text{for } \mathrm{P}_n\text{-almost all } x \in \mathbb{R}^n,$$

where $Q(j)$ denotes the cone of all points $x \in \mathbb{R}^n$ such that $x_j \geqslant \max_{i \neq j} x_i$. We can approximate the indicator function of $Q(j)$ by a smooth function to see that $M \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ and $D_j M = \mathbb{1}_{Q(j)}$ a.s. for all $1 \leqslant j \leqslant n$. Let

$$J(x) := \arg\max(x).$$

Clearly, $J(x)$ is defined uniquely for P_n-almost every $x \in \mathbb{R}^n$. For all other values of x, redefine $J(x) := 0$ to be concrete. Our computation of $D_j M$ equivalently yields

$$(DM)(x) = \mathbf{e}_{J(x)} \qquad \text{for } \mathrm{P}_n\text{-almost all } x \in \mathbb{R}^n, \tag{2.2}$$ eq:DM

where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ denote the standard basis of $\mathbb{R}^n$.

Let us end this section by introducing a little more notation.

The preceding discussion constructs, for every function $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$, the Malliavin derivative Df as an $\mathbb{R}^n$-valued function with coordinates in $L^2(\mathrm{P}_n)$. We will use the following natural notations exchangeably:

$$(Df)(x\,, j) := [(Df)(x)]_j = (D_j f)(x),$$

for every $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$, $x \in \mathbb{R}^n$, and $1 \leqslant j \leqslant n$. In this way we may also think of Df as a scalar-valued element of the real Hilbert space $L^2(\mathrm{P}_n \times \chi_n)$, where

Definition 1.9. χ_n always denotes the counting measure on $\{1\,, \ldots, n\}$.

We see also that the inner product on $\mathbb{D}^{1,2}(\mathrm{P}_n)$ is

$$\begin{aligned} \langle f\,, g\rangle_{1,2} &= \langle f\,, g\rangle_{L^2(\mathrm{P}_n)} + \langle Df\,, Dg\rangle_{L^2(\mathrm{P}_n \times \chi_n)} \\ &= \mathrm{E}(fg) + \mathrm{E}\,(Df \cdot Dg) \qquad\qquad \text{for all } f, g \in \mathbb{D}^{1,2}(\mathrm{P}_n). \end{aligned}$$

Definition 1.10. The random variable $Df \in L^2(\mathrm{P}_n \times \chi_n)$ is called the *Malliavin derivative* of the random variable $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$.

2 Higher-Order Derivatives

One can define higher-order weak derivatives just as easily as we obtained the directional weak derivatives.

Choose and fix $f \in C^2(\mathbb{R}^n)$ and two integers $1 \leqslant i, j \leqslant n$. The *mixed derivative* of f in direction $(i\,, j)$ is the function $x \mapsto (\partial^2_{i,j} f)(x)$, where

$$\partial^2_{i,j} f := \partial_i \partial_j f = \partial_j \partial_i f.$$

The *Hessian operator* ∇^2 is defined as

$$\nabla^2 := \begin{pmatrix} \partial^2_{1,1} & \cdots & \partial^2_{1,n} \\ \vdots & \ddots & \vdots \\ \partial^2_{n,1} & \cdots & \partial^2_{n,n} \end{pmatrix}.$$

With this in mind, we can define a Hilbertian inner product $\langle \cdot\,, \cdot \rangle_{2,2}$ via

$$\begin{aligned} \langle f\,, g\rangle_{2,2} &:= \int fg\,\mathrm{dP}_n + \int (\nabla f)\cdot(\nabla g)\,\mathrm{P}_n(\mathrm{d}x) + \int \mathrm{tr}\left[(\nabla^2 f)(\nabla^2 g)\right]\mathrm{dP}_n \\ &= \int f(x)g(x)\,\mathrm{P}_n(\mathrm{d}x) + \sum_{i=1}^n \int (\partial_i f)(x)(\partial_i g)(x)\,\mathrm{P}_n(\mathrm{d}x) \\ &\qquad\qquad + \sum_{i,j=1}^n \int (\partial^2_{i,j} f)(x)(\partial^2_{i,j} g)(x)\,\mathrm{P}_n(\mathrm{d}x) \\ &= \langle f\,, g\rangle_{1,2} + \int (\nabla^2 f)\cdot(\nabla^2 g)\,\mathrm{dP}_n \\ &= \mathrm{E}(fg) + \mathrm{E}\left[\nabla f\cdot\nabla g\right] + \mathrm{E}\left[\nabla^2 f\cdot\nabla^2 g\right] \qquad [f, g \in C^2_0(\mathrm{P}_n)], \end{aligned}$$

where $K\cdot M$ denotes the matrix—or Hilbert–Schmidt—inner product,

$$K\cdot M := \sum_{i,j=1}^n K_{i,j}M_{i,j} = \mathrm{tr}(K'M),$$

for all $n\times n$ matrices K and M.

We also obtain the corresponding Hilbertian norm $\|\cdot\|_{2,2}$ where:

$$\begin{aligned} \|f\|^2_{2,2} &= \|f\|^2_{L^2(\mathrm{P}_n)} + \sum_{i=1}^n \|\partial_i f\|^2_{L^2(\mathrm{P}_n)} + \sum_{i,j=1}^n \left\|\partial^2_{i,j} f\right\|^2_{L^2(\mathrm{P}_n)} \\ &= \|f\|^2_{1,2} + \left\|\nabla^2 f\right\|^2_{L^2(\mathrm{P}_n\times\chi^2_n)} \\ &= \mathrm{E}\left(f^2\right) + \mathrm{E}\left(\|\nabla f\|^2\right) + \mathrm{E}\left(\|\nabla^2 f\|^2\right) \qquad [f \in C^2_0(\mathrm{P}_n)]; \end{aligned}$$

$\chi^2_n := \chi_n \times \chi_n$ denotes the counting measure on $\{1\,, \cdots, n\}^2$; and

$$\|K\| := \sqrt{K\cdot K} = \sqrt{\sum_{i,j=1}^n K^2_{i,j}} = \sqrt{\mathrm{tr}(K'K)}$$

denotes the Hilbert–Schmidt norm of any $n\times n$ matrix K.

Definition 2.1. The *Gaussian Sobolev space* $\mathbb{D}^{2,2}(\mathrm{P}_n)$ is the completion of $C^2_0(\mathrm{P}_n)$ in the norm $\|\cdot\|_{2,2}$.

For every $f \in \mathbb{D}^{2,2}(\mathrm{P}_n)$ we can find functions $f_1, f_2, \ldots \in C_0^2(\mathrm{P}_n)$ such that $\|f_\ell - f\|_{2,2} \to 0$ as $\ell \to \infty$ Then $D_i f$ and $D^2_{i,j} f := \lim_{\ell\to\infty} \partial^2_{i,j} f$ exist in $L^2(\mathrm{P}_n)$ for every $1 \leqslant i, j \leqslant n$. Equivalently, $Df = \lim_{\ell\to\infty} \nabla f$ exists in $L^2(\mathrm{P}_n \times \chi_n)$ and $D^2 f = \lim_{\ell\to\infty} \nabla^2 f$ exists in $L^2(\mathrm{P}_n \times \chi_n^2)$.

Now we extend the definition to derivatives of order greater than two. Choose and fix an integer $k \geqslant 2$. If $q = (q_1, \ldots, q_k)$ is a vector of k integers in $\{1, \ldots, n\}$, then

$$(\partial^k_q f)(x) := (\partial_{q_1} \cdots \partial_{q_k} f)(x) \qquad [f \in C^k(\mathbb{R}^n),\, x \in \mathbb{R}^n].$$

Let ∇^k denote the formal k-tensor whose q-th coordinate is ∂^k_q. We can define a Hilbertian inner product $\langle \cdot , \cdot \rangle_{k,2}$ inductively via

$$\langle f , g\rangle_{k,2} = \langle f , g\rangle_{k-1,2} + \int (\nabla^k f) \cdot (\nabla^k g)\, \mathrm{dP}_n,$$

for all $f, g \in C_0^k(\mathrm{P}_n)$, where "$\cdot$" denotes the Hilbert–Schmidt inner product for k-tensors:

$$K \cdot M := \sum_{q\in\{1,\ldots,n\}^k} K_q M_q,$$

for all k-tensors K and M. The corresponding norm is defined via $\|f\|_{k,2} := \langle f , f\rangle^{1/2}_{k,2}$.

Definition 2.2. The *Gaussian Sobolev space* $\mathbb{D}^{k,2}(\mathrm{P}_n)$ is the completion of $C_0^k(\mathrm{P}_n)$ in the norm $\|\cdot\|_{k,2}$. We also define $\mathbb{D}^{\infty,2}(\mathrm{P}_n) := \cap_{k\geqslant 1}\mathbb{D}^{k,2}(\mathrm{P}_n)$.

If $f \in \mathbb{D}^{k,2}(\mathrm{P}_n)$ then we can find a sequence of functions $f_1, f_2, \ldots \in C_0^k(\mathrm{P}_n)$ such that $\|f_\ell - f\|_{k,2} \to 0$ as $\ell \to \infty$. It then follows that

$$D^j f := \lim_{\ell\to\infty} \nabla^j f_\ell \quad \text{exists in } L^2(\mathrm{P}_n \times \chi_n^j),$$

for every $1 \leqslant j \leqslant k$, where $\chi_n^j := \chi_n \times \cdots \times \chi_n$ [$j-1$ times] denotes the counting measure on $\{1, \ldots, n\}^j$. The operator D^k is called the kth *Malliavin derivative*.

It is easy to see that the Gaussian Sobolev spaces are nested; that is,

$$\mathbb{D}^{k,2}(\mathrm{P}_n) \subset \mathbb{D}^{k-1,2}(\mathrm{P}_n) \qquad \text{for all } 2 \leqslant k \leqslant \infty.$$

Also, whenever $f \in C_0^k(\mathrm{P}_n)$, the kth Malliavin derivative of f is just the classically-defined derivative $\nabla^k f$, which is a k-dimensional tensor. Because every polynomial in n variables is in $C_0^\infty(\mathrm{P}_n)$[3], it follows immediately that $\mathbb{D}^{\infty,2}(\mathbb{R}^n)$ contains all n-variable polynomials; and that all Malliavin derivatives acts as one might expect them to. This last fact will be important for the *Wiener chaos decomposition*, which is a way to write a fairly generic random variable as an infinite sum of polynomials, much like a Taylor series does. If the required sum converges properly then the last fact says that the Malliavin derivative acts on it as we expect it should.

More generally, we have the following.

[3]A function $f : \mathbb{R}^n \to \mathbb{R}$ is a polynomial in n variables if it can be written as a linear combination of monomials $x_1^{d_1} \cdots x_n^{d_n}$, where each d_j is a non-negative integer. The *degree* of each monomial is the sum of the exponents appearing in it, and the degree of the polynomial is the maximum degree of all monomials appearing in it. Thus, for example $g(x) = x_1 x_2^3 - 2x_5$ is a polynomial of degree 4 in 5 variables.

?⟨def:D:k,p⟩? **Definition 2.3.** For every integer $k \geqslant 1$ and real $p \geqslant 1$, the Gaussian Sobolev spaces $\mathbb{D}^{k,p}(\mathrm{P}_n)$ is defined as the completion of the space $C_0^\infty(\mathrm{P}_n)$ in the norm

$$\|f\|_{\mathbb{D}^{k,p}(\mathrm{P}_n)} := \|f\|_{k,p} := \left[\|f\|_{L^p(\mathrm{P}_n)}^p + \sum_{j=1}^k \|D^j f\|_{L^p(\mathrm{P}_n\times\chi_n^j)}^p\right]^{1/p}.$$

Each $\mathbb{D}^{k,p}(\mathrm{P}_n)$ is a Banach space in the preceding norm. Note that, as usual, these norms are not induced by an inner product unless $p = 2$. Furthermore, for each fixed k the spaces $\mathbb{D}^{k,p}$ are non-increasing in p.

3 The Adjoint Operator

Recall the canonical Gaussian probability density function $\gamma_n := \mathrm{d}\mathrm{P}_n/\mathrm{d}x$ from (1.1). Since $(D_j\gamma_n)(x) = -x_j\gamma_n(x)$, we can apply integration by parts and the product rule to see that for every $f, g \in C_0^1(\mathrm{P}_n)$,

$$\begin{aligned}
\mathrm{E}\left[D_j(f)g\right] &= \int_{\mathbb{R}^n} (D_j f)(x)g(x)\gamma_n(x)\,dx \\
&= -\int_{\mathbb{R}^n} f(x)D_j\left[g(x)\gamma_n(x)\right]\mathrm{d}x \\
&= -\int_{\mathbb{R}^n} f(x)(D_j g)(x)\,\mathrm{P}_n(\mathrm{d}x) + \int_{\mathbb{R}^n} f(x)g(x)x_j\,\mathrm{P}_n(\mathrm{d}x),
\end{aligned}$$

for $1 \leqslant j \leqslant n$. Using the $L^2(\mathrm{P}_n)$ inner product notation we can rewrite the latter identity as the "adjoint relation,"

$$\mathrm{E}\left[D_j(f)g\right] = \langle D_j f\,, g\rangle_{L^2(\mathrm{P}_n)} = \langle f\,, A_j g\rangle_{L^2(\mathrm{P}_n)} = \mathrm{E}\left[fA_j(g)\right], \tag{2.3}$$ IbP

where A is the formal adjoint of D; that is,

$$(Ag)(x) := -(Dg)(x) + xg(x). \tag{2.4}$$ A:g

Note that $g : \mathbb{R}^n \to \mathbb{R}$ is a real-valued function, but $Ag : \mathbb{R}^n \to \mathbb{R}^n$, and

$$(A_j g)(x) = -(D_j g)(x) + x_j g(x).$$

Furthermore, (2.4) is defined pointwise whenever $g \in C_0^1(\mathrm{P}_n)$, but it also makes sense as an identity in $L^2(\mathrm{P}_n \times \chi_n)$ if, for example, $g \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ and $x \mapsto xg(x)$ is in $L^2(\mathrm{P}_n \times \chi_n)$.

Let us pause to emphasize that (2.3) can be stated equivalently as

$$\mathrm{E}[gD(f)] = \mathrm{E}[fA(g)], \tag{2.5}$$ D:delta

as n-vectors.[4]

If $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$, then we can always find functions $f_1, f_2, \ldots \in C_0^1(\mathrm{P}_n)$ such that $\|f_\ell - f\|_{1,2} \to 0$ as $\ell \to \infty$. Note that

$$\begin{aligned}
\left\|\int gDf_\ell\,\mathrm{d}\mathrm{P}_n - \int gDf\,\mathrm{d}\mathrm{P}_n\right\| &\leqslant \|g\|_{L^2(\mathrm{P}_n)}\|Df_\ell - Df\|_{L^2(\mathrm{P}_n\times\chi_n)} \\
&\leqslant \|g\|_{L^2(\mathrm{P}_n)}\|f_\ell - f\|_{1,2} \to 0,
\end{aligned} \tag{2.6}$$ DfDf

[4] If $\zeta = (\zeta_1\,, \ldots, \zeta_m)$ is a random m-vector then $\mathrm{E}(\zeta)$ is the m-vector whose jth coordinate is $\mathrm{E}(\zeta_j)$.

as $\ell \to \infty$. Also,

$$\left\| \int f_\ell Ag \, \mathrm{dP}_n - \int f Ag \, \mathrm{dP}_n \right\| \leqslant \|Ag\|_{L^2(\mathrm{P}_n \times \chi_n)} \|f_\ell - f\|_{L^2(\mathrm{P}_n)} \\ \leqslant \|Ag\|_{L^2(\mathrm{P}_n \times \chi_n)} \|f_\ell - f\|_{1,2} \to 0, \tag{2.7}$$ fDgfDg

whenever $g \in C_0^1(\mathrm{P}_n)$. We can therefore combine (2.5), (2.6), and (2.7) in order to see that (2.5) in fact holds for all $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ and $g \in C_0^1(\mathrm{P}_n)$.

Finally define

$$\mathrm{Dom}[A] := \left\{ g \in \mathbb{D}^{1,2}(\mathrm{P}_n) : \; Ag \in L^2(\mathrm{P}_n \times \chi_n) \right\}. \tag{2.8}$$?Dom:A?

Since $C_0^1(\mathrm{P}_n)$ is dense in $L^2(\mathrm{P}_n)$, we may infer from (2.5) and another density argument the following.

?⟨pr:adjoint⟩? **Proposition 3.1.** *The adjoint relation* (2.5) *is valid for all* $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ *and* $g \in \mathrm{Dom}[A]$*.*

Definition 3.2. The linear operator A is the *adjoint operator*, and Dom$[A]$ is called the *domain of the definition*—or just *domain*—of A.

The linear space Dom$[A]$ has a number of nicely-behaved subspaces. The following records an example of such a subspace.

⟨pr:Subspace⟩ **Proposition 3.3.** *For every* $2 < p \leqslant \infty$*,*

$$\mathbb{D}^{1,2}(\mathrm{P}_n) \cap L^p(\mathrm{P}_n) \subset \mathrm{Dom}[A].$$

Proof. We apply Hölder's inequality to see that

$$\mathrm{E}\left(\|Z\|^2 [g(Z)]^2\right) = \int \|x\|^2 [g(x)]^2 \, \mathrm{P}_n(\mathrm{d}x) \leqslant c_p \|g\|^2_{L^p(\mathrm{P}_n)},$$

where

$$c_p = \left[\mathrm{E}\left(\|Z\|^{2p/(p-2)} \right) \right]^{(p-2)/(2p)} < \infty.$$

Therefore, $Zg(Z) \in L^2(\mathrm{P}_n \times \chi_n)$, and we may apply (2.4) to find that

$$\|Ag\|_{L^2(\mathrm{P}_n \times \chi_n)} \leqslant \|Dg\|_{L^2(\mathrm{P}_n \times \chi_n)} + c_p^{1/2} \|g\|_{L^p(\mathrm{P}_n)} \leqslant \|g\|_{1,2} + c_p^{1/2} \|g\|_{L^p(\mathrm{P}_n)} < \infty.$$

This proves that $g \in \mathrm{Dom}[A]$. □

Problems

⟨pbm:Ck_moments⟩ 1. Prove Lemma 1.2.
2. For which values of $s \in \mathbb{R}$ is $m(s) := \mathrm{E}(\|Z\|^s)$ finite? When it is finite, compute $\mathrm{E}(\|Z\|^s)$ in terms of the gamma function. These constants arose earlier during the course of the proof of Proposition 3.3.
3. Let $Z : \mathbb{R}^n \to \mathbb{R}^n$ denote the usual vector of independent, standard normal random variables, and define $X := \|MZ\|^s$, where M is a nonrandom $n \times n$ matrix and $s > 0$ is a non random real number. For which values of s and k is $X \in \mathbb{D}^{k,2}(\mathbb{R}^n)$?

⟨pbm:C^k_0⟩ 4. Prove that $f \in C_0^k(\mathrm{P}_n)$ if and only if f is infinitely differentiable in all of its variables, and

$$\lim_{\|x\|\to\infty} \mathrm{e}^{-\varepsilon\|x\|^2}|f(x)| = \lim_{\|x\|\to\infty} \mathrm{e}^{-\varepsilon\|x\|^2}\left|(\partial_{i_1}\cdots\partial_{i_m}f)(x)\right| = 0,$$

for all $1 \leqslant i_1, \ldots, i_m \leqslant n$ and $1 \leqslant m \leqslant k$.

5. Show directly from integration by parts that the standard Laplace operator

$$\Delta := D \cdot D = \sum_{i=1}^n \partial^2_{i,i}$$

is not self-adjoint on $L^2(\mathrm{P}_n)$, even though it is self-adjoint on the Lebesgue space $L^2(\mathbb{R}^n)$. What is the adjoint of Δ on $L^2(\mathrm{P}_n)$?

⟨pbm:Malliavin:Sobolev⟩ 6. Let $C_c^\infty(\mathbb{R}^n)$ denote the collection of all infinitely-differentiable functions of compact support from $\mathbb{R}^n$ to $\mathbb{R}$, and recall that the Sobolev space $W^{1,2}(\mathbb{R}^n)$ is the completion of $C_c^\infty(\mathbb{R}^n)$ in the norm $\|f\|_{L^2(\mathbb{R}^n)} + \|\nabla f\|_{L^2(\mathbb{R}^n)}$ for every $f \in C_c^\infty(\mathbb{R}^n)$. Construct an element of $\mathbb{D}^{1,2}(\mathrm{P}_n)$ that is not an element of $W^{1,2}(\mathbb{R}^n)$.
7. Fill in the details of the derivation of the identity (2.2).
8. If $G = (g_1, \ldots, g_n)$, then define $\delta G := A \cdot G = \sum_{i=1}^n A_i g_i$, when possible.
   (a) Verify that if every g_i is in $C_0^1(\mathrm{P}_n)$, then

   $$(\delta G)(x) = -(\operatorname{div} G)(x) + x \cdot G(x) \qquad \text{in } L^2(\mathrm{P}_n).$$

   (b) Prove the following *integration by parts formula*,

   $$\mathrm{E}\left[G \cdot (Df)\right] = \mathrm{E}\left[\delta(G)f\right],$$

   for all $f \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ and all random variables $G : \mathbb{R}^n \to \mathbb{R}^n$ that satisfy $\delta G \in L^2(\mathrm{P}_n)$.
9. Define $N := \max_{1\leqslant j\leqslant n} |Z_j|$. Prove that $N \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ and evaluate DN.

?⟨pbm:MG:transform⟩? 10. Suppose $n \geqslant 2$ is an integer. A stochastic process $X_1, \ldots, X_n$ is said to be *adapted* if X_i is measurable with respect to the σ-algebra generated by $Z_1, \ldots, Z_i$ for every $i = 1, \ldots, n$. Given an adapted process X, define a new stochastic process M – a so-called *martingale transform* of Z – as follows:

$$M_0 := 0 \quad \text{and} \quad M_k := \sum_{i=2}^k X_{i-1}Z_i \quad \text{for } k = 2, \ldots, n.$$

Suppose $X_i \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ for every $i = 2, \ldots, n$.
   (a) Prove that M is a mean-zero martingale and $M_i \in \mathbb{D}^{1,2}(\mathrm{P}_n)$ for every $i = 1, \ldots, n$.

(b) Compute DM_i for all $i = 1, \ldots, n$.

11. The purpose of this problem is to define an extension of the Malliavin derivative to more general, "non canonical" Gaussian spaces. Another, more standard, extension can be found in XXXX.
Let $(T, \mathcal{T}, \mathrm{Q})$ be a probability space that is rich enough to support countably-many independent standard-normal random variables $X_1, X_2, \ldots$. Let us say that a random variable Y on $(T, \mathcal{T}, \mathrm{Q})$ is *smooth and cylindrical* [written $Y \in \mathscr{C}$] if there exists an integer $n \geqslant 1$ and a function $\varphi \in C_0^\infty(\mathrm{P}_n)$ such that $Y = \varphi(X_1, \ldots, X_n)$. For every smooth random variable Y of this form, define $\|Y\|_{1,2} := \|\varphi\|_{1,2}$, where the second norm is the one that was defined in this chapter. Let $\mathbb{D}^{1,2}(\mathrm{Q})$ denote the completion of $\mathscr{C}$ in the norm $\| \cdot \|_{1,2}$ thus defined.
    (a) If $Y \in \mathscr{C}$ has the form $Y = \varphi(X_1, \ldots, X_n)$ for some $\varphi \in C_0^\infty(\mathrm{P}_n)$, then define its Malliavin derivative as $DY := (\nabla\varphi)(X_1, \ldots, X_n)$. Prove that the linear operator D has a unique extension, which we continue to denote by D, to $\mathbb{D}^{1,2}(\mathrm{Q})$.
    (b) Prove that, for every $Y \in \mathbb{D}^{1,2}(\mathrm{Q})$, we can identify DY with a sequence $\{D_j Y\}_{j=1}^\infty$ of random variables and that $D_j X_i = 1$ if $i = j$ and 0 otherwise.
    (c) Let $\{a_i\}_{i=1}^\infty$ be a sequence of constants such that $\sum_{i=1}^\infty a_i^2 < \infty$.
        i. Verify that, if $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, then $Y := \sum_{i=1}^\infty a_i f(X_i)$ converges almost surely and in $L^2(\mathrm{Q})$.
        ii. Verify that $Y \in \mathbb{D}^{1,2}(\mathrm{Q})$ and that $D_j Y = a_j f'(X_j)$ almost surely for every $j \geqslant 1$.
    (d) Identify the adjoint A of D on $\mathscr{C}$ and write an integration by parts formula that generalizes the adjoint relation (2.5) that was valid on a canonical Gauss space.