Calculus in Gauss Space

1. The Gradient Operator

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

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

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,

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

and so on, notation being typically obvious from context.

Let $\partial_j := \partial/\partial x_j$ for all $1 \leqslant j \leqslant n$ and let $\nabla := (\partial_1, \ldots, \partial_n)$ denote the gradient operator, acting on continuously-differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$. From now on we will use the following.

Definition 1.1. Let $C_0^k(\mathbb{P}_n)$ 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 $\varepsilon > 0$. Equivalently, $f \in C_0^k(\mathbb{P}_n)$ 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 $1 \leqslant m \leqslant k$. We also define

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

We will frequently use the following without mention.

Lemma 1.2. *If $f \in C_0^k(\mathbb{P}_n)$, then*

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

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

I omit the proof since it is elementary.

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

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

Notice that $\|\cdot\|_{1,2}$ is a *bona fide* Hilbertian norm on $C_0^2(\mathbb{P}_n)$ with Hilbertian inner product

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

We will soon see that $C_0^2(\mathbb{P}_n)$ is not a Hilbert space with the preceding norm and inner product because it is not complete. This observation prompts the following definition.

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

In order to understand what the elements of $\mathbb{D}^{1,2}(\mathbb{P}_n)$ look like consider a function $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$. By definition we can find a sequence

$f_1, f_2, \ldots \in C_0^1(\mathbb{P}_n)$ such that $\|f_\ell - f\|_{1,2} \to 0$ as $\ell \to \infty$. Since $L^2(\mathbb{P}_n)$ is complete, we can deduce also that

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

It follows, by virtue of construction, that

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

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

In general, we can think of elements of $\mathbb{D}^{1,2}(\mathbb{P}_n)$ as functions in $L^2(\mathbb{P}_n)$ that have one weak derivative in $L^2(\mathbb{P}_n)$. 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 $\mathbb{D}^{1,2}(\mathbb{P}_n)$ was, but the Lebesgue measure is used in place of $\mathbb{P}_n$ everywhere. Since $\gamma_n(x) = \mathrm{d}\mathbb{P}_n(x)/\mathrm{d}x < 1$,[1] it follows that the Hilbert space $\mathbb{D}^{1,2}(\mathbb{P}_n)$ is richer than the Hilbert space $W^{1,2}(\mathbb{R}^n)$, whence the Malliavin derivative is an extension of Sobolev's [generalized] gradient.[2]

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

ex:Smoothing:1 **Example 1.4** ($n = 1$). Consider the case $n = 1$ and let

$$f(x) := (1 - |x|)_+ \qquad \text{for all } x \in \mathbb{R}.$$

Then we claim that $f \in \mathbb{D}^{1,2}(\mathbb{P}_1) \setminus C_0^1(\mathbb{P}_1)$ and in fact we have the $\mathbb{P}_1$-a.s. identity,[3]

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

whose intuitive meaning ought to be clear.

In order to prove these assertions let ψ_1 be a symmetric probability density function on $\mathbb{R}$ such that $\psi_1 \in C^\infty(\mathbb{R})$, $\psi_1 \equiv$ a positive constant on $[-1\,,1]$, and $\psi_1 \equiv 0$ off of $[-2\,,2]$. 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

---

[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] The extension is strict. For instance, $f(x) := \exp(x)$ $[x \in \mathbb{R}]$ defines a function in $\mathbb{D}^{1,2}(\mathbb{P}_1) \setminus W^{1,2}(\mathbb{R})$.

[3] It might help to recall that Df is defined as an element of the Hilbert space $L^2(\mathbb{P}_1)$ 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 $\mathbb{P}_1$-a.s. equality is the same thing as Lebesgue-a.e. equality.

$N \to \infty$ because f is uniformly continuous. In particular, $\|f_N - f\|_{L^2(\mathbb{P}_n)} \to 0$ as $N \to \infty$. To complete the proof it remains to verify that

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

By the dominated convergence theorem and integration by parts,

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

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

We can apply a change of variables, together with the symmetry of ψ_1, in order to see that $A_N(x) = \int_{-Nx}^{N(1-x)} \psi_1(y)\,\mathrm{d}y$. Therefore, $A_N(x) \to -\mathrm{sign}(x)\mathbb{1}_{[0,\infty)}(x)$ as $N \to \infty$ for all $x \neq 0$. Since $A_N(x) + \mathrm{sign}(x)\mathbb{1}_{[0,\infty)}(x)$ is bounded uniformly by 2, the dominated convergence theorem implies that $A_N(x) \to -\mathrm{sign}(x)\mathbb{1}_{[-1,1]}(x)$ as $N \to \infty$ in $L^2(\mathbb{P}_1)$. 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 $F \in \mathbb{D}^{1,2}(\mathbb{P}_n) \setminus C_0^1(\mathbb{P}_n)$ 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 $\mathbb{D}^{1,2}(\mathbb{P}_n)$, $F_N \in C_0^1(\mathbb{P}_n)$, and $F \notin C_0^1(\mathbb{P}_n)$. Thus, it follows that $F \in \mathbb{D}^{1,2}(\mathbb{P}_n) \setminus C_0^1(\mathbb{P}_n)$. 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 $\mathbb{P}_n$-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, $\mathbb{P}_n$-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, $\mathbb{E}(|f|^k) < \infty$ for all $k \geqslant 1$. A density argument, similar to the one that appeared in the preceding examples, shows that $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$ and

$$\|Df\| \leqslant \text{Lip}(f) \qquad \text{a.s.}$$

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

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) to wit. The following ought to reinforce this point of view.

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

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

Proof. If f and ψ are continuously differentiable 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. In general we appeal to 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

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

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 $M \in \mathbb{D}^{1,2}(\mathbb{P}_n)$ 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 is defined uniquely for $\mathbb{P}_n$-almost every $x \in \mathbb{R}^n$. And our computation of $D_j M$ is equivalent to

$$(DM)(x) = {}_{J(x)} \qquad \text{for } \mathbb{P}_n\text{-almost all } x \in \mathbb{R}^n,$$

where ${}_1, \ldots, {}_n$ 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 $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$, the Malliavin derivative Df as an $\mathbb{R}^n$-valued function with coordinates in $L^2(\mathbb{P}_n)$. We will use the following natural notations exchangeably:

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

for every $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$, $x \in \mathbb{R}^n$, 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 $L^2(\mathbb{P}_n \times \chi_n)$, where

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

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

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

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

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

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

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

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

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:

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

$\chi^2_n := \chi_n\times\chi_n$ 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* $\mathbb{D}^{2,2}(\mathbb{P}_n)$ is the completion of $C^2_0(\mathbb{P}_n)$ in the norm $\|\cdot\|_{2,2}$.

For every $f\in\mathbb{D}^{2,2}(\mathbb{P}_n)$ we can find functions $f_1,f_2,\ldots\in C^2_0(\mathbb{P}_n)$ 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 $L^2(\mathbb{P}_n)$ for every $1\leqslant i,j\leqslant n$. Equivalently, $Df = \lim_{\ell\to\infty}\nabla f$ exists in $L^2(\mathbb{P}_n\times\chi_n)$ and $D^2 f = \lim_{\ell\to\infty}\nabla^2 f$ exists in $L^2(\mathbb{P}_n\times\chi^2_n)$.

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 write

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

Let ∇^k denote the formal k-tensor whose q-th coordinate is ∂_q^k.

We define a Hilbertian inner product $\langle \cdot, \cdot \rangle_{k,2}$ [inductively] via

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

for all $f, g \in C_0^k(\mathbb{P}_n)$, where "$\cdot$" denotes the Hilbert–Schmidtt 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 $\|f\|_{k,2} := \langle f, f \rangle_{2,2}^{1/2}$.

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

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

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

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,

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

Also, whenever $f \in C_0^k(\mathbb{P}_n)$, 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 $C_0^\infty(\mathbb{P}_n)$,[4] 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.

More generally, we have the following.

[4]A function $f : \mathbb{R}^n \to \mathbb{R}$ is a polynomial in n variables if it can be written as $f(x) = f_1(x_1) \times \cdots \times f_n(x_n)$, for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, where each f_j is a polynomial on $\mathbb{R}$. The *degree* of the polynomial f is the maximum of the degrees of $f_1, \ldots, f_n$. Thus, for example $f(x) = x_1 x_2^3 - 2x_5$ is a polynomial of degree 3 in 5 variables.

def:D:k,p **Definition 2.3.** We define Gaussian Sobolev spaces $\mathbb{D}^{k,p}(\mathbb{P}_n)$ by completing the space $C_0^\infty(\mathbb{P}_n)$ in the norm

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

Each $\mathbb{D}^{k,p}(\mathbb{P}_n)$ is a Banach space in the preceding norm.

3. The Adjoint Operator

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

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

for $1 \leqslant j \leqslant n$. Let $g := (g_1, \ldots, g_n)$ and sum the preceding over all $1 \leqslant j \leqslant n$ to find the following "adjoint relation,"

$$\mathbb{E}\left[D_j(f)g\right] = \langle D_j f, g\rangle_{L^2(\mathbb{P}_n)} = \langle f, A_j g\rangle_{L^2(\mathbb{P}_n)} = \mathbb{E}\left[fA_j(g)\right], \tag{2.2}$$ IbP

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

$$(Ag)(x) := -(Dg)(x) + xg(x). \tag{2.3}$$ A:g

Eq. (2.3) is defined pointwise whenever $g \in C_0^1(\mathbb{P}_n)$. But it also makes sense as an identity in $L^2(\mathbb{P}_n \times \chi_n)$ if, for example, $g \in \mathbb{D}^{1,2}(\mathbb{P}_n)$ and $x \mapsto xg(x)$ is in $L^2(\mathbb{P}_n \times \chi_n)$.

Let us pause to emphasize that (2.2) can be stated equivalently as

$$\mathbb{E}[gD(f)] = \mathbb{E}[fA(g)], \tag{2.4}$$ D:delta

as n-vectors.[5]

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

$$\begin{aligned}\left\|\int gDf_\ell\,\mathrm{d}\mathbb{P}_n - \int g\cdot Df\,\mathrm{d}\mathbb{P}_n\right\| &\leqslant \|g\|_{L^2(\mathbb{P}_n\times\chi_n)}\|Df_\ell - Df\|_{L^2(\mathbb{P}_n\times\chi_n)}\\ &\leqslant \|g\|_{L^2(\mathbb{P}_n\times\chi_n)}\|f_\ell - f\|_{2,1} \to 0,\end{aligned} \tag{2.5}$$ DfDf

[5] If $W = (W_1, \ldots, W_m)$ is a random m-vector then $\mathbb{E}(W)$ is the m-vector whose jth coordinate is $\mathbb{E}(W_j)$.

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

$$\left\| \int f_\ell Ag \, \mathrm{d}\mathbb{P}_n - \int fAg \, \mathrm{d}\mathbb{P}_n \right\| \leqslant \|Ag\|_{L^2(\mathbb{P}_n)} \|f_\ell - f\|_{L^2(\mathbb{P}_n)} \\ \leqslant \|Ag\|_{L^2(\mathbb{P}_n)} \|f_\ell - f\|_{1,2} \to 0, \tag{2.6}$$

fDgfDg

whenever $g \in C_0^1(\mathbb{P}_n)$. We can therefore combine (2.4), (2.5), and (2.6) in order to see that (2.4) in fact holds for all $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$ and $g \in C_0^1(\mathbb{P}_n)$.

Finally define

$$\mathrm{Dom}[A] := \left\{ g \in \mathbb{D}^{1,2}(\mathbb{P}_n) : \; Ag \in L^2(\mathbb{P}_n \times \chi_n) \right\}. \tag{2.7}$$

Dom:A

Since $C_0^1(\mathbb{P}_n)$ is dense in $L^2(\mathbb{P}_n)$, we may infer from (2.4) and another density argument the following.

pr:adjoint

Proposition 3.1. *The adjoint relation* (2.4) *is valid for all* $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$ *and* $g \in \mathrm{Dom}[A]$.

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

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

pr:Subspace

Proposition 3.3. *For every* $2 < p \leqslant \infty$,

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

Proof. We apply Hölder's inequality to see that

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

where

$$c_p = \left[\mathbb{E}\left(\|Z\|^{2p/(p-1)} \right) \right]^{(p-1)/(2p)} = \left[\int \|x\|^{2p/(p-1)} \, \mathbb{P}_n(\mathrm{d}x) \right]^{(p-1)/(2p)} < \infty.$$

Therefore, $Zg(Z) \in L^2(\mathbb{P}_n \times \chi_n)$, and we may apply (2.3) to find that

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

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

Very often, people prefer to use the *divergence operator* δ associated to A in place of A itself. That is, if $G = (g_1, \ldots, g_n)$ with every g_i in the "domain of A_i," then

$$(\delta G)(x) := (A \cdot G)(x) := \sum_{i=1}^{n} (A_i g_i)(x).$$

If every g_i is in $C_0^1(\mathbb{P}_n)$, then (2.3) shows that

$$(\delta G)(x) = -\sum_{i=1}^{n}(D_i g_i)(x) + \sum_{i=1}^{n} x_i g_i(x) = -(\operatorname{div} g)(x) + x \cdot g(x),$$

where "div" denotes the usual divergence operator in Lebesgue space.

We will be working directly with the adjoint in these notes, and will keep the discussion limited to A, rather than δ. Still, it is worth mentioning the scalar identity,

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

for all functions $G : \mathbb{R}^n \to \mathbb{R}^n$ for which δG can be defined in $L^2(\mathbb{P}_n)$ and all $f \in \mathbb{D}^{1,2}(\mathbb{P}_n)$. The preceding formula is aptly known as the *integration by parts formula* of Malliavin calculus, and is equivalent to the statement that A and D are $L^2(\mathbb{P}_n)$-adjoints of one another, though one has to pay attention to the domains of the definition of A, D, and δ carefully in order to make precise this equivalence.