Notes on Malliavin calculus 1

Dario Gasbarra

4. helmikuuta 2011

Gaussian integration by parts

Lemma

Let $G(\omega)$ a real valued gaussian random variable with $E(G)=0$ and variance $E(G^2)=\sigma^2$, If f, h are smooth functions, such that $f, h, f', h' \in L^2(\mathbb{R},\gamma)$

$$E(f'(G)h(G)) = E\left(f(G)\left(\frac{h(G)G}{E(G^2)} - h'(G) \right) \right) \qquad (0.1)$$

Proof

$P(G \in dx) = \gamma(x)dx$ with density

$$\gamma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) .$$

Note that

$$\frac{d}{dx}\gamma(x) = -\frac{\gamma(x)x}{\sigma^2}$$

Proof

$P(G \in dx) = \gamma(x)dx$ with density

$$\gamma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) .$$

Note that

$$\frac{d}{dx}\gamma(x) = -\frac{\gamma(x)x}{\sigma^2}$$

Integrating by parts

$$\int_{-\infty}^{\infty} f'(x)h(x)\gamma(x)dx = -\int_{-\infty}^{\infty} f(x)\frac{d}{dx}\Big(h(x)\gamma(x)\Big)dx$$

$$= \int_{-\infty}^{\infty} f(x)\left(\frac{h(x)x}{\sigma^2} - h'(x)\right)\gamma(x)dx \; \square$$

Denote

$$\partial f(x) := f'(x) \text{ and } \partial^* h(x) := \left(\frac{h(x)x}{\sigma^2} - h'(x) \right)$$

Denote

$$\partial f(x) := f'(x) \text{ and } \partial^* h(x) := \left(\frac{h(x)x}{\sigma^2} - h'(x) \right)$$

Then

$$(\partial f, h)_{L^2(\mathbb{R},\gamma)} = (f, \partial^* h)_{L^2(\mathbb{R},\gamma)}$$

∂^* is the adjoint of the derivative operator in $L^2(\mathbb{R}, \gamma)$.

Denote

$$\partial f(x) := f'(x) \text{ and } \partial^* h(x) := \left(\frac{h(x)x}{\sigma^2} - h'(x) \right)$$

Then

$$(\partial f, h)_{L^2(\mathbb{R},\gamma)} = (f, \partial^* h)_{L^2(\mathbb{R},\gamma)}$$

∂^* is the adjoint of the derivative operator in $L^2(\mathbb{R}, \gamma)$.

Definition

We say that $f \in L^2(\mathbb{R}, \gamma)$, has weak derivative $g \in L^2(\mathbb{R}, \gamma)$ in Sobolev sense if $\forall h$ with classical derivative h' such that $\partial^* h \in L^2(\gamma)$,

$$\int_{\mathbb{R}} g(x)h(x)\gamma(x)dx = \int_{\mathbb{R}} f(x)\partial^* h(x)\gamma(x)dx$$

and we denote $\partial f = f' := g$.

This definition extends the classical derivative. We introduce the weighted Sobolev space

$$W^{1,2}(\mathbb{R},\gamma) := \{\ f \in L^2 : f \text{ Sobolev differentiable } \}$$

with norm

$$\| f \|^2_{W^{1,2}(\gamma)} = \| f \|^2_{L^2(\gamma)} + \| \partial f \|^2_{L^2(\gamma)}$$

Proposition

The set of smooth functions with derivatives of polynomial growth is dense in $W^{1,2}(\gamma)$

Proposition

The gaussian integration by parts formula

$$E_P(\partial f(G)h(G)) = E_P(f(G)\partial^* h(G))$$

extends to $f \in W^{1,2}(\mathbb{R},\gamma)$ $h \in Domain(\partial^*)$.

Corollary

For $h(x) \equiv 1$, $f \in W^{1,2}(\mathbb{R},\gamma)$

$$E(f'(G)) = \frac{E(f(G)G)}{E(G^2)}$$

Linear regression

Let $X(\omega), Y(\omega) \in L^2(P)$. Then

$$\widehat{X}(\omega) = \widehat{b} + \widehat{a} Y(\omega) \quad \text{with} \tag{0.2}$$

$$\widehat{a} = \frac{E(X(Y - E(Y)))}{E(Y^2) - E(Y)^2} \tag{0.3}$$

$$\widehat{b} = E(X) - \widehat{b} E(Y) \tag{0.4}$$

Linear regression

Let $X(\omega), Y(\omega) \in L^2(P)$. Then

$$\widehat{X}(\omega) = \widehat{b} + \widehat{a} Y(\omega) \quad \text{with} \tag{0.2}$$

$$\widehat{a} = \frac{E(X(Y - E(Y)))}{E(Y^2) - E(Y)^2} \tag{0.3}$$

$$\widehat{b} = E(X) - \widehat{b} E(Y) \tag{0.4}$$

is the L^2-projection of X on the linear subspace generated by Y, such that

$$E((\widehat{X} - X)^2) = \min_{a,b \in \mathbb{R}} E\Big((a + bY - X)^2 \Big)$$

In general $\widehat{X}(\omega) \neq E(X|\sigma(Y))(\omega)$, which is the projection of X on the subspace $L^2(\Omega, \sigma(Y), P)$.

In general $\widehat{X}(\omega) \neq E(X|\sigma(Y))(\omega)$, which is the projection of X on the subspace $L^2(\Omega, \sigma(Y), P)$.
When (X, Y) is jointly gaussian, $\widehat{X} = E(X|\sigma(Y))$.

In general $\widehat{X}(\omega) \neq E(X|\sigma(Y))(\omega)$, which is the projection of X on the subspace $L^2(\Omega, \sigma(Y), P)$.
When (X, Y) is jointly gaussian, $\widehat{X} = E(X|\sigma(Y))$.
Let $G \sim \mathcal{N}(0, \sigma^2)$ and consider $F = f(G)$ for some non-linear function $f \in W^{1,2}(\mathbb{R}, \gamma)$. By 0.2 the best linear estimator of $f(G)$ given G is

$$\widehat{f(G)} = E(f(G)) + \frac{E(f(G)G)}{E(G^2)} G$$

In general $\widehat{X}(\omega) \neq E(X|\sigma(Y))(\omega)$, which is the projection of X on the subspace $L^2(\Omega, \sigma(Y), P)$.
When (X, Y) is jointly gaussian, $\widehat{X} = E(X|\sigma(Y))$.
Let $G \sim \mathcal{N}(0, \sigma^2)$ and consider $F = f(G)$ for some non-linear function $f \in W^{1,2}(\mathbb{R}, \gamma)$. By 0.2 the best linear estimator of $f(G)$ given G is

$$\begin{aligned} \widehat{f(G)} &= E(f(G)) + \frac{E(f(G)G)}{E(G^2)} G \\ &= E(f(G)) + E(f'(G))G \quad \text{by 0.1} \end{aligned}$$

We have

$$f(G) = E_P(f(G)|\sigma(G)) = E(f(G)) + E(f'(G))G + M^f \quad (0.5)$$

Clearly $E(M^f) = 0$, but also

$$E(M^f G) = E\left(\{ f(G) - E(f(G)) - E(f'(G))G \} G \right)$$

$$= E\left(\left\{ f(G) - E(f(G)) - \frac{E(f(G)G)}{E(G^2)} G \right\} G \right) = 0$$

Clearly $E(M^f) = 0$, but also

$$E(M^f G) = E\Big(\{ f(G) - E(f(G)) - E(f'(G))G \} G \Big)$$

$$= E\left(\left\{ f(G) - E(f(G)) - \frac{E(f(G)G)}{E(G^2)} G \right\} G \right) = 0$$

The linearization error M^f is uncorrelated with G.

Lemma

Assume that f', f'' are bounded and continuous. Then

$$\frac{E((M^f)^2)}{\sigma^2} \to 0 \quad \text{as } \sigma \to 0$$

Proof. Let $G(\omega) \sim \mathcal{N}(0,1)$, obviously $\sigma G(\omega) \sim \mathcal{N}(0,\sigma^2)$

$$f(\sigma G) = f(0) + f'(0)\sigma G + \frac{1}{2} f''(X_\sigma)\sigma^2 G^2$$

with $X_\sigma \to 0$ a.s. as $\sigma \to 0$

$$E\left(\left\{E(f(\sigma G)) + E(f'(\sigma G))\sigma G - f(\sigma G)\right\}^2\right) =$$

$$\mathrm{Var}\left(\{\, f'(0) - E(f'(\sigma G))\}\sigma G + \frac{1}{2} f''(X_\sigma)\sigma^2 G^2\right)$$

$$= \{\, f'(0) - E(f'(\sigma G))\}^2\sigma^2 + \frac{\sigma^4}{4}\mathrm{Var}(f''(X_\sigma)G^2)$$

$$+\{\, f'(0) - E(f'(\sigma G))\}\sigma^3 E(f''(X_\sigma)G^3)$$

Divide by σ^2 and let $\sigma \to 0$. Since the derivatives are bounded and continuos, the result follows by Lebesgue dominated convergence $\square$

Let $\Delta W_1, \dots, \Delta W_n$ i.i.d. Gaussian with $E(\Delta W_1) = 0$ $E(\Delta W_1^2) = \sigma^2$. These are consecutive increments of the random walk $W_m = \sum_{k=1}^{m} \Delta W_k$.
Let

$$F(\omega) = f(\Delta W_1(\omega), \dots, \Delta W_n(\omega))$$

with $f(x_1, \dots, x_n) \in W^{1,2}(\mathbb{R}^n, \gamma^{\otimes n})$.
Let $\mathcal{F}_k = \sigma(\Delta W_1, \dots, \Delta W_k)$

Let $\Delta W_1, \dots, \Delta W_n$ i.i.d. Gaussian with $E(\Delta W_1) = 0$ $E(\Delta W_1^2) = \sigma^2$. These are consecutive increments of the random walk $W_m = \sum_{k=1}^m \Delta W_k$.
Let

$$F(\omega) = f(\Delta W_1(\omega), \dots, \Delta W_n(\omega))$$

with $f(x_1, \dots, x_n) \in W^{1,2}(\mathbb{R}^n, \gamma^{\otimes n})$.
Let $\mathcal{F}_k = \sigma(\Delta W_1, \dots, \Delta W_k)$

Lemma

We have the martingale representation

$$F = E(F) + \sum_{k=1}^n E(\partial_k f(\Delta W_1, \dots, \Delta W_n) | \mathcal{F}_{k-1}) \Delta W_k + M_n$$

where M is a $(\mathcal{F}_k)$-martingale with $M_0 = 0$ and $\langle M, W \rangle = 0$.

By induction it is enough to show that

$$E(F|\mathcal{F}_k) = \\ E(F|\mathcal{F}_{k-1}) + E(\partial_k f(\Delta W_1, \dots, \Delta W_n)|\mathcal{F}_{k-1})\Delta W_k + \Delta M_k$$

with

$$E(\Delta M_k|\mathcal{F}_{k-1}) = 0, \ E(\Delta W_k \Delta M_k|\mathcal{F}_{k-1}) = 0 \qquad (0.6)$$

Let's fix k and consider the enlarged σ-algebra

$$\mathcal{G}_{k-1} = \sigma(\Delta W_1, \dots, \Delta W_{k-1}, \Delta W_{k+1}, \dots \Delta W_n) \supseteq \mathcal{F}_{k-1}$$

By fixing $(\Delta W_i, i \neq k)$ applying 0.5 to the k-the coordinate ΔW_k

$$F = E(F|\mathcal{G}_{k-1}) + E(\partial_k(\Delta W_1, \ldots \Delta W_k)|\mathcal{G}_{k-1})\Delta W_k + \Delta \widetilde{M}_k$$

By the indepedence of the increments

$$\begin{aligned}
& f(\Delta W_1, \ldots \Delta W_n) \\
& = E\Big(f(x_1, \ldots, x_{k-1}, \Delta W_k, x_{k+1}, \ldots x_n) \Big)\Big|_{x_i = \Delta W_i,\ i \neq k} \\
& + E\Big(\partial_k f(x_1, \ldots, x_{k-1}, \Delta W_k, x_{k+1}, \ldots x_n) \Big)\Big|_{x_i = \Delta W_i,\ i \neq k} \Delta W_k \\
& + \Delta \widetilde{M}_k
\end{aligned}$$

By fixing $(\Delta W_i, i \neq k)$ applying 0.5 to the k-the coordinate ΔW_k

$$F = E(F|\mathcal{G}_{k-1}) + E(\partial_k(\Delta W_1, \dots \Delta W_k)|\mathcal{G}_{k-1})\Delta W_k + \Delta \widetilde{M}_k$$

By the indepedence of the increments

$$\begin{aligned} & f(\Delta W_1, \dots \Delta W_n) \\ & = E\left(f(x_1, \dots, x_{k-1}, \Delta W_k, x_{k+1}, \dots x_n)\right)\Big|_{x_i = \Delta W_i,\ i\neq k} \\ & + E\left(\partial_k f(x_1, \dots, x_{k-1}, \Delta W_k, x_{k+1}, \dots x_n)\right)\Big|_{x_i = \Delta W_i,\ i\neq k} \Delta W_k \\ & + \Delta \widetilde{M}_k \end{aligned}$$

with

$$E(\Delta \widetilde{M}_k|\mathcal{G}_{k-1}) = 0, \quad E(\Delta \widetilde{M}_k \Delta W_k|\mathcal{G}_{k-1}) = 0,$$

which implies

$$E(\Delta \widetilde{M}_k|\mathcal{F}_{k-1}) = 0, \quad E(\Delta \widetilde{M}_k \Delta W_k|\mathcal{F}_{k-1}) = 0.$$

By taking conditional expectation w.r.t. $\mathcal{F}_k$ and using independence of increments

$$E(F|\mathcal{F}_k) =$$
$$E\left(f\left(x_1, \dots, x_{k-1}, \Delta W_k, \Delta W_{k+1}, \dots \Delta W_n\right)\right)\Big|_{x_i=\Delta W_i,\, i<k} +$$
$$E\left(\partial_k f\left(x_1, \dots, x_{k-1}, \Delta W_k, \Delta W_{k+1}, \dots \Delta W_n\right)\right)\Big|_{x_i=\Delta W_i,\, i<k} \Delta W_k$$
$$+\Delta M_k$$

where

$$\Delta M_k := E(\Delta \widetilde{M}_k|\mathcal{F}_k)$$

with

$$E(\Delta M_k|\mathcal{F}_{k-1}) = E(\Delta M_k \Delta W_k|\mathcal{F}_{k-1}) = 0 \quad \Box$$

Assuming that the derivatives $\partial_k f, \partial_k^2 f$ are bounded and continuous, by Jensen inequality and lemma 2

$$\sigma^{-2} E((\Delta M_k)^2 | \mathcal{F}_{k-1}) \leq \sigma^{-2} E((\Delta \widetilde{M}_k)^2 | \mathcal{F}_{k-1}) \longrightarrow 0$$

P a.s. as $\sigma \longrightarrow 0$, and by dominated convegence

$$\sigma^{-2} E((\Delta M_k)^2) \leq \sigma^{-2} E((\Delta \widetilde{M}_k)^2) \longrightarrow 0$$

Definition

The (finite-dimensional) Malliavin derivative is the random gradient

$$DF := \nabla f(\Delta W_1, \dots, \Delta W_n) \in \mathbb{R}^n$$

Brownian motion, $(W_t : t \in [0, T])$ is a gaussian process with $W_0 = 0$ and such that for every n, $0 = t_0 \le t_1 \le \cdots \le t_n = T$ the increments $(W_{t_i} - W_{t_{i-1}})$ are independent and gaussian with variances $(t_i - t_{i-1})$.
We will show that there is realization of Brownian motion as a random continuous function.
Suppose that we have a random variable $F(\omega)$ which is measurable with respect to the σ-algebra $\mathcal{F}_t^W = \sigma(W_s : 0 \le s \le t)$, and that this can be approximated by random variables of the form

$$F_n(\omega) := f_n\Big(W_{t_1^{(n)}} - W_{t_0^{(n)}}, \dots, W_{t_n^{(n)}} - W_{t_{n-1}^{(n)}}\Big)$$

with smooth f_n and $t_k^{(n)} := Tk/n$.

Under some regularity assumptions as $n \to \infty$ the orthogonal linearization error in

$$F_n = E(F_n) + \sum_{k=1}^{n} E(\nabla_k F_n | \mathcal{F}^{(n)}_{k-1}) \Delta W^{(n)}_k + M^n_n$$

vanishes (in $L^2(P)$ sense) and the limit is the Clark-Ocone martingale representation

$$F = E(F) + \int_0^T E(D_s F | \mathcal{F}^W_s) dW_s$$

where the Ito integral appears.

Skorokhod integral

In $L^2(\mathbb{R}^n, \gamma^{\otimes n}(x)dx)$ the Malliavin derivative of
$F = f(\Delta W_1, \dots, \Delta W_n)$ as the random gradient
$DF = \nabla f(\Delta W_1, \dots, \Delta W_n)$, where ΔW_k are i.i.d. $\mathcal{N}(0, \Delta t)$
Let $u_k = u_k(\Delta W_1, \dots, \Delta W_n)$ for $k = 1, \dots, n$.
Introduce the scalar product

$$\langle u, v \rangle := \Delta t \sum_{k=1}^{n} u_k v_k$$

We give the n-dimensional generalization of the 1-dimensional integration by parts formula.
We need a random variable which we denote by $\delta(u)$ (the Skorokhod integral or divergence integral) such that

$$E(\langle DF, u\rangle) = E(F\delta(u))$$

for all smooth random variables F.

Rewrite the left hand side

$$\Delta t \sum_{k=1}^{n} E\big(u_k(\Delta W_1,\dots,\Delta W_n)\partial_k f(\Delta W_1,\dots,\Delta W_n)\big)$$

by independence and the 1-dimensional gaussian integration by parts

$$= \Delta t \sum_{k=1}^{n} E(\partial_k^* u_k(\Delta W_1,\dots,\Delta W_n) f(\Delta W_1,\dots,\Delta W_n))$$

$$= E\left(F \Delta t \left(\sum_{k=1}^{n} \frac{u_k \Delta W_k}{\Delta t} - \sum_{k=1}^{n} \partial_k u_k \right)\right)$$

Rewrite the left hand side

$$\Delta t \sum_{k=1}^{n} E\big(u_k(\Delta W_1, \dots, \Delta W_n)\partial_k f(\Delta W_1, \dots, \Delta W_n)\big)$$

by independence and the 1-dimensional gaussian integration by parts

$$= \Delta t \sum_{k=1}^{n} E(\partial_k^* u_k(\Delta W_1, \dots, \Delta W_n) f(\Delta W_1, \dots, \Delta W_n))$$

$$= E\left(F \Delta t \left(\sum_{k=1}^{n} \frac{u_k \Delta W_k}{\Delta t} - \sum_{k=1}^{n} \partial_k u_k \right) \right)$$

so that

$$\delta(u) = \sum_{k=1}^{n} u_k \Delta W_k - \sum_{k=1}^{n} D_k u_k \Delta t$$

The first term is a Riemann sum, while the second term is called Malliavin trace.

When $u_k = u_k(\Delta W_1, \ldots, \Delta W_{k-1}, \Delta W_{k+1}, \ldots, \Delta W_n)$ does not depend on ΔW_k , the Malliavin trace vanishes.

When $u_k = u_k(\Delta W_1, \dots, \Delta W_{k-1}, \Delta W_{k+1}, \dots, \Delta W_n)$ does not depend on ΔW_k , the Malliavin trace vanishes.
For $F \equiv 1$, $DF \equiv 0$ and when exists $\delta(u) \in L^2(\Omega)$, necessarily

$$E(\delta(u)) = E(\langle u, 0 \rangle) = 0$$

When $u_k = u_k(\Delta W_1, \dots, \Delta W_{k-1}, \Delta W_{k+1}, \dots, \Delta W_n)$ does not depend on ΔW_k , the Malliavin trace vanishes.
For $F \equiv 1$, $DF \equiv 0$ and when exists $\delta(u) \in L^2(\Omega)$, necessarily

$$E(\delta(u)) = E(\langle u, 0 \rangle) = 0$$

In the continuous time case the Skorokhod integral is given by

$$\delta(u) := \int_0^T u_s \delta W_s = \int_0^T u_s dW_s - \int_0^T D_s u_s ds$$

where $\int_0^T u_s dW_s$ is a backward integral defined as the limit in probability or $L^2(P)$-sense of the Riemann sums, and the last term is the Malliavin trace.
When u is adapted, that is u is $\mathcal{F}_s^W$-measurable for all s the Malliavin trace vanishes and the Skorokhod integral coincides with the Ito integral.

Note that if φ is smooth, $D\varphi(F) = \varphi'(F)DF$. We have also the product rule $D(FG) = G\ DF + F\ DG$.
Consider a process $u_k = u_k(\Delta W_1, \dots, \Delta W_n)$

$$E\left(\delta\left(\frac{u}{\langle u, DF\rangle}\right)\varphi(F)\right) = E\left(\left\langle \frac{u}{\langle u, DF\rangle}, D\varphi(F)\right\rangle\right)$$
$$= E\left(\frac{\varphi'(F)}{\langle u, DF\rangle}\langle u, DF\rangle\right) = E\big(\varphi'(F)\big)$$

This holds for all choices of (u_k) and φ. By taking $u = DF$ we obtain

$$E\big(\varphi'(F)\big) = E\left(\varphi(F)\delta\left(\frac{DF}{\parallel DF \parallel^2}\right)\right)$$

where

$$\parallel DF \parallel^2 = \langle DF, DF\rangle = \Delta t \sum_k^n (D_k F)^2$$

Computation of densities

Let $F = f(\Delta W_1, \dots, \Delta W_n)$ a random variable with Malliavin Sobolev derivative. For $a < b \in \mathbb{R}$ consider

$$\psi(x) = \int_a^b \mathbf{1}(r \leq x) dr$$

which is continuous and piecewise linear with Sobolev derivative $\psi'(x) = \mathbf{1}_{[a,b]}(x)$.

$$P(a < F \leq b) = \int_a^b p_F(r) dr \text{ (when } F \text{ has density)}$$

$$= E_P\big(\mathbf{1}(a < F \leq b)\big) = E_P\big(\psi'(F)\big) = E_P\left(\psi(F)\delta\left(\frac{DF}{\| DF \|^2}\right)\right)$$

$$= E_P\left(\delta\left(\frac{DF}{\| DF \|^2}\right) \int_a^b \mathbf{1}(r \leq F) dr\right) = \text{ (Fubini)}$$

$$= \int_a^b E_P\left(\mathbf{1}(r \leq F)\delta\left(\frac{DF}{\| DF \|^2}\right)\right) dr$$

This implies

$$p_F(r) = E_P\left(\mathbf{1}(r \le F)\delta\left(\frac{DF}{\| DF \|^2}\right)\right) = E_P\left(\mathbf{1}(r \le F)Y\right)$$

with Malliavin weight

$$Y := \delta\left(\frac{DF}{\| DF \|^2}\right) =$$

$$\frac{1}{\| DF \|^2}\sum_{k=1}^n D_k F \Delta W_k - \sum_{k=1}^n D_k\left(\frac{D_k F}{\| DF \|^2}\right)\Delta t$$

$$= \frac{1}{\| DF \|^2}\sum_{k=1}^n D_k F \Delta W_k - \frac{1}{\| DF \|^2}\sum_{k=1}^n D^2_{kk} F \Delta t$$

$$+ \frac{2}{\| DF \|^4}\sum_{k=1}^n \sum_{h=1}^n D_k F \; D_h F \; D^2_{kh} F \; \Delta t$$

For $F = f(\Delta W_1, \dots, \Delta W_n)$ we need that $f \in C^2$ and integrability conditions. The formula extends to the infinite-dimensional case when F is a smooth functional of the Brownian path.

For $F = f(\Delta W_1, \dots, \Delta W_n)$ we need that $f \in C^2$ and integrability conditions. The formula extends to the infinite-dimensional case when F is a smooth functional of the Brownian path.
For $i \in \mathbb{N}$ let $(\Delta W_1^{(i)}, \dots, \Delta W_n^{(i)})$, i.i.d copies of the gaussian vector, let

$$
\begin{aligned}
F^{(i)} &:= f(\Delta W_1^{(i)}, \dots, \Delta W_n^{(i)}), \\
Y^{(i)} &:= Y(\Delta W_1^{(i)}, \dots, \Delta W_n^{(i)})
\end{aligned}
$$

We estimate $p_F(t)$ by Monte Carlo

$$
\widehat{p}_F^{(M)}(r) = \frac{1}{M} \sum_{i=1}^{M} Y^{(i)} \mathbf{1}(F^{(i)} \geq r)
$$

There are other choices for the Malliavin weight: for

$$u_k = \frac{1}{n\Delta t \; D_k F}$$

we obtain

$$E(\langle u, D\varphi(F)\rangle) = \frac{1}{n\Delta t} E\big(\varphi'(F)DF, \langle (DF)^{-1}\rangle\big) =$$
$$= \frac{1}{n\Delta t} E\left(\varphi'(F)\sum_{k=1}^{n}(D_k F)^{-1} D_k F \Delta t\right)$$
$$= E(\varphi'(F)) = E\big(\varphi(F)U\big)$$

with Malliavin weight

$$U = \frac{1}{n\Delta t}\delta\big((DF)^{-1}\big) \tag{0.7}$$

Example: quadratic functional

Let

$$F = \sum_{k=1}^{n} W_k^2 \Delta t \text{ with}$$

$$D_h F = 2 \sum_{k=h}^{n} W_k \Delta t, \quad D_{h,h}^2 F = 2(n - h + 1)\Delta t$$

We compute the Malliavin weight 0.7

$$U = \frac{1}{n\Delta t}\left(\sum_{h=1}^{n} \frac{1}{D_h F} dW_h - \sum_{h=1}^{n} D_h\big((D_h F)^{-1}\big)\Delta t\right)$$

$$= \frac{1}{2n\Delta t} \sum_{h=1}^{n} \left(\sum_{k=h}^{n} W_k \Delta t\right)^{-1} \Delta W_h$$

$$+ \frac{1}{n\Delta t} \sum_{h=1}^{n} \left(2 \sum_{k=h}^{n} W_k \Delta t\right)^{-2} 2(n-h+1)(\Delta t)^2 =$$

$$\frac{1}{2n(\Delta t)^2} \left\{ \sum_{h=1}^{n} \left(\sum_{k=h}^{n} W_k\right)^{-1} \Delta W_h \right.$$

$$\left. + \sum_{h=1}^{n} \left(\sum_{k=h}^{n} W_k\right)^{-2} (n-h+1)\Delta t \right\}$$

Counterexample: Maximum of gaussian random walk

Let $W_0 = 0$, $W_m = \sum_{k=1}^{m} \Delta W_k$ for $m = 1, \dots, n$ the gaussian random walk, and let

$$F = W_n^* := \max_{m=0,1,\dots,n} \{ W_m \} = f(\Delta W_1, \dots, \Delta W_n)$$

Let

$$\tau_n = \tau_n(W_1, \dots, W_n) = \arg \max_{m=0,1,\dots,n} W_m$$

the random time where the maximum is achieved. Note that with positive probability $W_n^* = 0$ and $\tau_n = 0$ when the random walk stays on the negative side, so we know that there is point mass at 0, W_n^* does not have a density.

Clearly for $k = 1, \dots, n$

$$D_k W_n^* = \partial_k f_n(\Delta W_1, \dots, \Delta W_n) = \mathbf{1}(\tau_n \geq k) \quad \text{a.s.}$$

The problem is that the indicator of a set is never Malliavin differentiable. In one dimension,let $\psi(x) = \mathbf{1}(x \geq r)$. The derivative $\psi'(x) = \delta_r(t)$ is the Dirac delta function of distribution theory (not to be confused with the Skorkohod integral $\delta(u)$!):

$$E(\psi'(x)h(x)) = E(\psi(x)\partial^* h(x))$$
$$\int_{-\infty}^{\infty} \psi(x)\partial^* h(x)\gamma(x)dx$$
$$-\int_r^{\infty} \frac{d}{dx}\Big(h(x)\gamma(x)\Big)dx = h(r)\gamma(r) - h(\infty)\gamma(\infty)$$
$$= h(r)\gamma(r) = \int_{\mathbb{R}} \delta_r(x)h(x)\gamma(x)dx$$

Now the Dirac $\delta_r(x)$ is a generalized function which is not in $L^2(\mathbb{R},\gamma)$. Therefore ψ does not have a Malliavin/Sobolev derivative.
The second order Malliavin derivative $D^2_{hk}X^*_n = D_h\mathbf{1}(\tau_n > k)$ doesn't exist as random variables in L^2 and the Malliavin weights are not well defined. In the so called white noise theory generalized random variables (also called random distributons) are introduced exactly in the same way as the generalized functions (also called distributions) of analysis.

Let $\gamma(x)$ be the standard gaussian density in $\mathbb{R}$.

Lemma

The polynomials are dense in $L^2(\mathbb{R}, \gamma)$.

Proof Otherwise there is a random variable $F = f(G) \in L^2(P)$ with $E(f(G)G^n) = 0$ $\forall n \in \mathbb{N}$ where G is standard gaussian. Consider the (signed) measure on $\mathbb{R}$

$$\mu(A) := E_P\big(f(G)\mathbf{1}_A(G)\big)$$

We show that $\mu \equiv 0$ which implies $f(G) = 0$ P a.s.
The Fourier transform of μ is

$$\widehat{\mu}(t) := E_P(f(G)\exp(itG))$$

For $t = (\sigma + \tau i) \in \mathbb{C}$ with $\sigma, \tau \in \mathbb{R}$,

$$\widehat{\mu}(t) := E_P\big(f(G)\exp(i\sigma G)\exp(-\tau G)\big)$$

Since

$$\begin{aligned}
& E_P\left(\left|\frac{\partial}{\partial\sigma}\left\{f(G)\exp(-\tau G)\exp(i\sigma G)\right\}\right|\right) \\
&= E_P\big(\big|f(G)\exp(-\tau G) iG \exp(i\sigma G)\big|\big) \\
&\le E_P\big(|f(G) G \exp(-\tau G)|\big) \\
&\le E_P\big(|f(G) G(\exp(-aG) + \exp(-bG))|\big)
\end{aligned}$$

where $\exp(-\tau G) \le \exp(-aG) + \exp(-bG)$ $\forall \tau \in (a, b) \subseteq \mathbb{R}$.

By Cauchy-Schwartz inequality

$$\leq E_P(f(G)^2)^{1/2} E(G^2\{\exp(-aG) + \exp(-bG))^2)^{1/2}$$
$$= E_P(f(G)^2)^{1/2}\{E(G^2 \exp(-2aG)) +$$
$$+E(G^2 \exp(-2bG)) + 2E(G^2 \exp(-(a+b)G))\}^{1/2} < \infty$$

by Lebesgue's dominated convergence theorem we can change the order of derivation and integration (Theorem A 16.1 in Williams' book)

$$\frac{\partial}{\partial\sigma}\widehat{\mu}(\tau + i\sigma) = i\ E_P\big(f(G) G \exp(i\sigma G) \exp(-\tau G)\big)$$

Similarly

$$\frac{\partial}{\partial\tau}\widehat{\mu}(\tau + i\sigma) = -E_P\big(f(G) G \exp(i\sigma G) \exp(-\tau G)\big) = i\ \frac{\partial}{\partial\sigma}\widehat{\mu}(\tau + i\sigma)$$

$\widehat{\mu} : \mathbb{C} \to \mathbb{C}$ is analytic since satisfies the Cauchy-Riemann condition.

Therefore has the power series expansion

$$\widehat{\mu}(t) = \sum_{t=0}^{\infty} \widehat{\mu}^{(n)}(0)\frac{t^n}{n!}$$

$$\mu^{(n)}(t) = \frac{d^n}{dt^n}\widehat{\mu}(t) = i^n E_P(f(G)\exp(itG)G^n),$$

$$\widehat{\mu}^{(n)}(0) = i^n E_P(f(G)G^n) = 0 \;\forall n \in \mathbb{N}$$

where by adapting the previous argument we can take derivatives inside the expectation. Therefore $\widehat{\mu}(t) = 0$ and by Lévy inversion theorem $\mu(dx) = 0$, which implies $E_P(f(G)^2) = 0$ $\square$.

Hermite polynomials in $L^2(\mathbb{R},\gamma)$.

Let G be a standard gaussian random variable with density $\gamma(x)$.
Define the (unnormalized) Hermite polynomials

$$h_0(x) \equiv 1,\ h_n(x) = (\partial^* h_{n-1})(x) = (\partial^{*n} 1)(x)$$

We have

- $$\partial h_n(x) = n h_{n-1}(x)$$

Hermite polynomials in $L^2(\mathbb{R},\gamma)$.

Let G be a standard gaussian random variable with density $\gamma(x)$.

Define the (unnormalized) Hermite polynomials

$$h_0(x) \equiv 1, \; h_n(x) = (\partial^* h_{n-1})(x) = (\partial^{*n} 1)(x)$$

We have

- $$\partial h_n(x) = n h_{n-1}(x)$$

- $$h_n(x) = \exp(x^2/2) \frac{d^n}{dx^n} \exp(-x^2/2)$$

  Ex: $h_1(x) = x$, $h_2(x) = (x^2 - 1)$, $h_3(x) = (x^3 - 3x)$, $h_4(x) = x^4 - 6x^2 + 3$, $h_5(x) = (x^5 - 10x^3 + 15x)$

Hermite polynomials in $L^2(\mathbb{R}, \gamma)$.

Let G be a standard gaussian random variable with density $\gamma(x)$.
Define the (unnormalized) Hermite polynomials

$$h_0(x) \equiv 1, \; h_n(x) = (\partial^* h_{n-1})(x) = (\partial^{*n} 1)(x)$$

We have

- $$\partial h_n(x) = n h_{n-1}(x)$$

- $$h_n(x) = \exp(x^2/2) \frac{d^n}{dx^n} \exp(-x^2/2)$$

  Ex: $h_1(x) = x, \; h_2(x) = (x^2 - 1), \; h_3(x) = (x^3 - 3x),$
  $h_4(x) = x^4 - 6x^2 + 3, \; h_5(x) = (x^5 - 10x^3 + 15x)$

- $$E_P(h_n(G) h_m(G)) = \delta_{n,m} n!$$

Since the polynomials are dense in $L^2(\mathbb{R},\gamma)$, it follows that the normalized Hermite polynomials

$$H_n(x) := \frac{h_n(x)}{\sqrt{n!}} \; n \in \mathbb{N}$$

form an orthonormal basis in $L^2(\mathbb{R},\gamma)$: for $f(G) \in L^2(P)$,

$$f(G) = \sum_{n=0}^{\infty} E_P(f(G)H_n(G))H_n(G) = \sum_{n=0}^{\infty} E_P(f(G)h_n(G))\frac{h_n(G)}{n!}$$

and when $f(x)$ is infinitely differentiable in Sobolev sense

$$= \sum_{n=0}^{\infty} E_P(f(G)(\partial^{*n}1)(G))\frac{h_n(G)}{n!} = \sum_{n=0}^{\infty} E_P(\partial^n f(G))\frac{h_n(G)}{n!} \quad (0.8)$$

the convergence is in $L^2(P)$ sense

$$E_P\left(\left\{f(G)-\sum_{n=1}^{M} E_P(f(G)H_n(G))H_n(G)\right\}^2\right)\to 0 \quad \text{as } M\uparrow\infty$$

Define the generating function

$$f(t,x):=\exp\big(tx-t^2/2\big)=\frac{\gamma(x-t)}{\gamma(x)}=\frac{d\mathcal{N}(t,1)}{d\mathcal{N}(0,1)}(x)$$

which is the density ratio for the gaussian shift $G\to(t+G)$
Note that $E_P(f(t,G))=1$. Since $f(t,x)\in C^\infty$, by 0.8

$$\exp\big(tx-t^2/2\big)=\sum_{n=0}^{\infty}E_P\left(\frac{d^n}{dx^n}f(t,G)\right)\frac{h_n(x)}{n!}$$

$$=\sum_{n=0}^{\infty}E_P\big(t^n f(t,G)\big)\frac{h_n(x)}{n!}=\sum_{n=0}^{\infty}h_n(x)\frac{t^n}{n!}$$

Note that

$$t^n = E_P\left(h_n(G)\exp\big(tG - t^2/2\big)\right) = E_P\big(h_n(t+G)\big)$$

where on the right side we have changed the measure.

Hermite polynomials in $L^2(\mathbb{R}^n, \gamma^{\otimes n})$.

Let $G = (G_1, \dots, G_n)$ a random vector with indepedent standard gaussian coordinates.
Since $L^2(\mathbb{R}^n, \gamma^{\otimes n}) = \overline{L^2(\mathbb{R}, \gamma)^n}$, which is the L^2-closure of the linear space containing the products $f_1(x_1)f_2(x_1)\dots f_n(x_n)$ with $f_i \in L^2(\mathbb{R}, \gamma)$,
the polynomials in the variables $x_1, \dots, x_n$ are dense in $L^2(\mathbb{R}^n, \gamma^{\otimes n})$.

Definition

$\alpha = (\alpha_1, \dots, \alpha_n)$ *with* $\alpha_i \in \mathbb{N}$ *is a multi-index.*
$\alpha! := \prod_{i=1}^n \alpha_i!$

For $x = (x_1, \ldots, x_n)$ define the unnormalized and normalized multivariate Hermite polynomials

$$h_\alpha(x) = \prod_{i=1}^n h_{\alpha_i}(x_i)$$

$$H_\alpha(x) = \prod_{i=1}^n H_{\alpha_i}(x_i) = \prod_{i=1}^n \frac{h_{\alpha_i}(x)}{\sqrt{\alpha_i!}} = \frac{h_\alpha(x)}{\sqrt{\alpha!}}$$

Lemma

$\{H_\alpha(x) : \alpha$ multi-index$\}$ is an orthonormal basis in $L^2(\mathbb{R}^n, \gamma^{\otimes n})$

Proof Let $\beta=(\beta_1,\dots,\beta_n)$ $\beta_i\in\mathbb{N}$,

$$E_P(H_\alpha(G)H_\beta(G)) = E_P\left(\prod_{i=1}^n H_{\alpha_i}(G_i)\prod_{j=1}^n H_{\beta_j}(G_j)\right) =$$

$$\prod_{i=1}^n E_P\big(H_{\alpha_i}(G_i)H_{\beta_i}(G_i)\big) = \prod_{i=1}^n \delta_{\alpha_i,\beta_i} = \delta_{\alpha,\beta}$$

Infinite dimensional gaussian space

$L^2(\mathbb{R}^{\mathbb{N}}, \gamma^{\otimes \mathbb{N}})$ is the space of sequences $x = (x_i : i \in \mathbb{N})$.
On this space we use the product σ-algebra $\mathcal{B}(\mathbb{R}^{\mathbb{N}}) = \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}$ which is the smallest σ-algebra such that the coordinate evaluations $x \mapsto x_i$ are measurable.
The Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is the smallest σ algebra containing the open sets.
The product measure $\gamma^{\otimes \mathbb{N}}$ is such that $\forall n \in \mathbb{N}$, $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$

$$\gamma^{\otimes \mathbb{N}}(\{ \ x : x_1 \in B_1, \dots, x_n \in B_n\}) = \prod_{i=1}^{n} \gamma(B_i)$$

Definition

$\alpha = (\alpha_i : i \in \mathbb{N})$ with $\alpha_i \in \mathbb{N}$ and

$$|\alpha| := \sum_{i=1}^{\infty} \alpha_i < \infty$$

is a multi-index

Definition

A polynomial in the variables $(x_i : i \in \mathbb{N})$ is given by

$$p(x) = c_0 + \sum_{i=1}^{\infty} c_i x_i^{\alpha_i}$$

$c_i \in \mathbb{R}$, and α is a multiindex, $|\alpha| < \infty$, which depends on finitely many coordinates.

Next we show

$$L^2(\mathbb{R}^{\mathbb{N}}, \gamma^{\otimes \mathbb{N}}) = \bigoplus_{n \in \mathbb{N}} L^2(\mathbb{R}^n, \gamma^{\otimes n})$$

It follows also that when $G = (G_i : i \in \mathbb{N})$ is a sequence of independent standard gaussian r.v.

$$\left\{ H_\alpha(G) := \prod_{i=1}^{\infty} H_{\alpha_i}(G_i), \ \alpha \text{ multindex }, |\alpha| < \infty \right\}$$

is an orthonormal basis.