

Derivada de Malliavin

Ahora que establecimos que las direcciones de Cameron-Martin nos permiten cambiar las integrales bajo medidas gaussianas, estamos listos para definir la derivada de Malliavin. Podemos hacerlo desde dos enfoques:

- Podemos hablar de caos polinómico y luego llegar a la derivada
- Definimos la derivada y no hacemos referencia alguna al caos polinómico

Seguiremos el segundo enfoque por razones pedagógicas, recopilando todo lo que hemos aprendido. Después de eso, seguiremos a Friz (2002) y Alos (2021) para ponerlo en terminos más concretos.

Fundamentos de la derivada de Malliavin

Hemos acumulado una gran cantidad de conocimiento. Es hora de que abordemos el objetivo de este ensayo. Cuando planteamos la Pregunta, mencionamos que la derivada de una variable aleatoria que sigue un movimiento browniano con respecto al tiempo no se puede calcular, ese era todo el punto: ¿Qué tipo de derivada podemos calcular?

Consideremos una dirección de Cameron-Martin de $\hat{h}(\tau) = \int_0^\tau h dt$. Además, $\hat{h} \in \hat{H}$, el espacio de Cameron-Martin. Esta definición nos permite tener una derivada en el tiempo para la dirección, por lo que $\frac{\partial \hat{h}(\tau)}{\partial \tau} = h$, con $h \in H = L^2[0, 1]$, el espacio de funciones que son continuas entre 0 y 1, y que permiten una norma-2¹.

Las trayectorias de movimiento browniano hasta τ se denotarán como $\omega(\tau)$. La función $1_{[0, \tau]}$ es una función que es 1 en el intervalo indicado y 0 en el resto del intervalo, comportándose

¹En términos más formales, podemos considerar esta traslación:

$$T_{\hat{h}}(\omega) = \omega + \varepsilon \int_0^\tau h = \omega + \varepsilon \hat{h}$$

Y el cambio de medida/variable, es decir, nuestra derivada de Radon-Nikodym, será:

$$\frac{d(T_{\hat{h}})_* W}{dW}(\omega) = \exp \left(\int_0^1 h d\beta(\omega) - \frac{1}{2} \int_0^1 h^2 d\lambda \right)$$

Finalmente, sin pérdida de generalidad, podemos suponer que $\|h\| = 1$ y que todos los h_i sean ortonormales en H . Esta es seguro de suponer debido a la ortonormalización de Gram-Schmidt.

como un filtro. Finalmente, definiremos una función como $W(1_{[0,\tau]})$ como una integral de Ito del filtro. También utilizaremos B_t y dB_t para hacer referencia a los movimientos brownianos, junto con W_t y dW_t .

Nuestro último ingrediente es una función polinómica² f que toma n movimientos brownianos como variables, cada uno deteniéndose en un momento diferente. Dos movimientos brownianos desde 0 hasta t son esencialmente iguales, incluso si sus trayectorias son diferentes. Así, cada movimiento browniano irá de 0 a t_1, t_2, \dots, t_n . Esta función f es por el momento bastante básica. Encontraremos su uso y extensiones más adelante.

Ahora, haremos algo como en la sección de cálculo de variaciones. Planteamos $f(\omega + \varepsilon \hat{h})$. Esto significa que estamos moviendo cada una de las n trayectorias del movimiento browniano un poco a lo largo de una dirección de Cameron-Martin, por lo que es una traslación aceptable. Finalmente, tomamos la derivada con respecto a ε , y evaluamos la derivada en $\varepsilon = 0$. Los movimientos brownianos, que no se pueden derivar, no se tocan porque no dependen de ε .

Y aquí está el resultado final:

$$\begin{aligned} \left. \frac{d}{d\varepsilon} F(\omega + \varepsilon \hat{h}) \right|_{\varepsilon=0} &= \left[\sum_{i=1}^n \partial_i f(\omega(t_1) + \varepsilon \hat{h}, \omega(t_2) + \varepsilon \hat{h}, \dots, \omega(t_n) + \varepsilon \hat{h}) \int_0^{t_i} h d\lambda \right]_{\varepsilon=0} \\ &= \sum_{i=1}^n \partial_i f(\omega(t_1), \omega(t_2), \dots, \omega(t_n)) \int_0^{t_i} h d\lambda \\ &= \langle DF, h \rangle_H \end{aligned}$$

Definimos DF como:

$$DF = \sum_i \partial_i f(W(1_{[0,\tau_1]}), \dots, W(1_{[0,\tau_n]})) 1_{[0,\tau_i]}$$

O su extensión alternativa:

$$DF = \sum_i \partial_i f(W(h_1), \dots, W(h_n)) h_i$$

haciendo que DF sea una variable aleatoria con valores de H .

²Una función con crecimiento polinómico también es aceptable. Las condiciones para esta función pertenecen a la teoría y no aportan claridad pedagógica.

¿Qué acaba de pasar?

Esta D es la derivada de Malliavin. Es un operador lineal (como los que vimos antes) y se aplica a la función F . Significa que si tenés una función F con variables compuestas por estos movimientos brownianos multiplicados por estas funciones h_i , entonces se puede aplicar esta definición de derivada.

Esta F parece muy limitada ahora mismo, pero si se pudiera *representar, aproximar o encontrar una equivalencia* entre la variable aleatoria de interés y una función como F , entonces se podría aplicar la derivada a la equivalencia. La teoría se ocupa de encontrar esas conexiones, no nos molestaremos en hacerlo aquí.

Llegados a este punto, la colección de funciones h_i son un requisito teórico y parecen salidas de la nada. Por un lado, no son más arbitrarias que el desplazamiento h de la definición clásica de derivada $\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$. En los textos que vienen, seleccionaremos las $h_i \in H$ que sean adecuadas para nuestros propósitos. Fuentes más teóricas pueden establecer que las h_i son ortonormales (o pueden volverse ortonormales), especialmente bajo un producto interno con un ponderador gaussiano. En vez de eso, tomamos la opción más directa y más simple: usaremos una única función $h = 1_{[0,1]}$, es decir, una función que es 1 en el intervalo de integración y 0 en el resto. Esto ayuda cuando la función está dentro de una integral porque “corta” o “filtra” el dominio de integración, como en este ejemplo:

$$\int_{-\infty}^{+\infty} f(x, t) \cdot 1_{[0,1]}(x) dx = \int_0^1 f(x, t) dx$$

Un segundo interrogante es si el intervalo $[0, 1]$ es un límite fundamental o si podemos extenderlo al menos hasta $[0, T]$, que es lo que vemos en procesos estocásticos comunes y corrientes. Todos los autores que se centran en los aspectos teóricos de la derivada se mantienen en ese intervalo, tratando aparentemente a 1 como un límite convencional. Por otro lado, Alos (2021) lo extiende a T . Por lo tanto, a partir de ahora usamos, basándonos puramente en la conveniencia, tanto $h = 1_{[0,T]}$ como $h = 1_{[0,1]}$.

Ejemplos

Ahora, comencemos con un ejemplo *muy* simple. Calculemos la derivada de Malliavin de un proceso de Wiener/movimiento browniano:

$$F = W(h) = \int_0^1 h dB$$

Sea h una función de L^2 en el intervalo $[0, 1]$, como hemos definido anteriormente. En este caso, la derivada de Malliavin es:

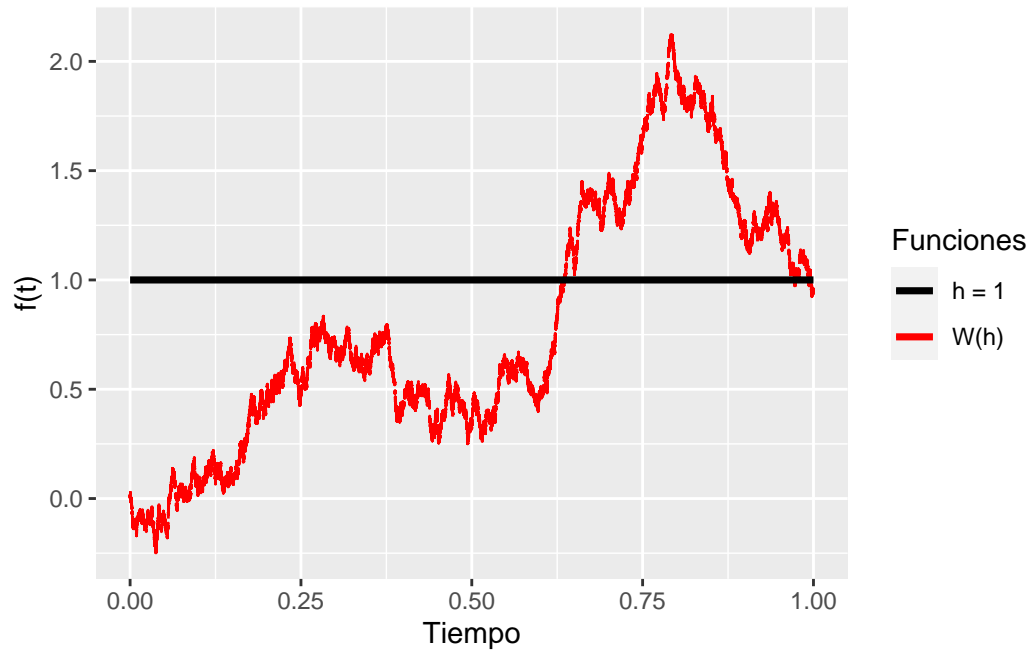
$$DF = D(W(h)) = h$$

Una buena regla mnemotécnica: esto es como si hiciéramos $\frac{\partial F}{\partial W} \cdot h$ en lugar de $\frac{\partial F}{\partial t}$ (algo que sabemos que no podemos hacer). Es decir, esto es cuánto cambia la función a medida que cambia el movimiento browniano subyacente (en la dirección de Cameron-Martin). Aquí hay un gráfico de cómo se ven las diferentes opciones de $h(t)$ y un ejemplo/realización de $W(h(t))$:

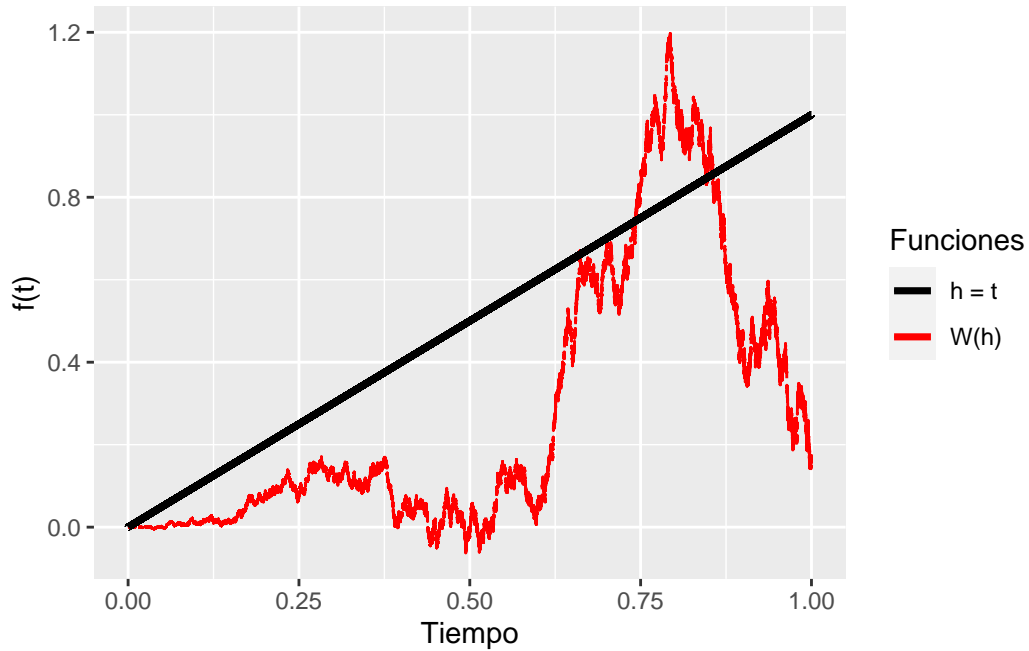
```
library(ggplot2)
library(rgl)

# Setup
steps = 100000
h <- 1/steps
t <- seq(from = 0, to = 1, length.out=steps)
dBt <- c(0, rnorm(n=(steps-1),mean = 0,sd = sqrt(h)))

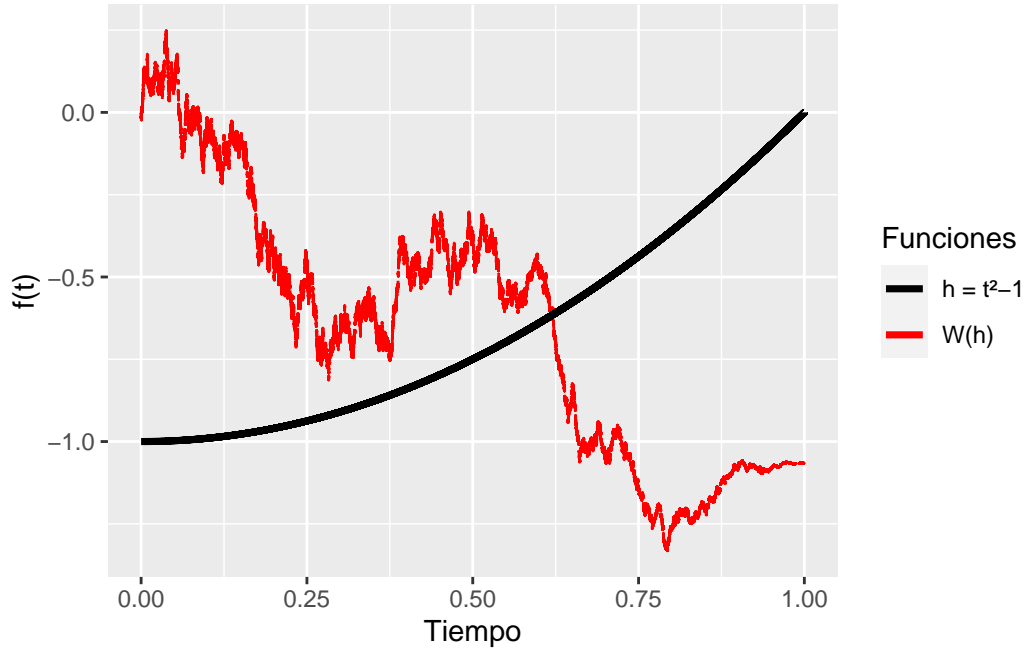
# Left
h_1 <- Vectorize(function(x) 1)
h_hat_1 <- cumsum(h_1(t) * dBt)
ggplot(mapping = aes(x = t)) +
  xlab('Tiempo') +
  ylab('f(t)') +
  geom_line(mapping = aes(y=h_hat_1, colour="W(h)"),linetype = 2) +
  geom_line(mapping = aes(y = h_1(t), colour = "h = 1"), linewidth = 1.25) +
  scale_colour_manual("Funciones",values=c("black","red"))
```



```
# Middle
h_2 <- Vectorize(function(x) x)
h_hat_2 <- cumsum(h_2(t) * dBt)
ggplot(mapping = aes(x = t)) +
  xlab('Tiempo') +
  ylab('f(t)') +
  geom_line(mapping = aes(y=h_hat_2, colour="W(h)"),linetype = 2) +
  geom_line(mapping = aes(y = h_2(t), colour = "h = t"), linewidth = 1.25) +
  scale_colour_manual("Funciones",values=c("black","red"))
```



```
# Right
h_3 <- Vectorize(function(x) x**2 - 1)
h_hat_3 <- cumsum(h_3(t) * dBt)
ggplot(mapping = aes(x = t)) +
  xlab('Tiempo') +
  ylab('f(t)') +
  geom_line(mapping = aes(y=h_hat_3, colour="W(h)"),linetype = 2) +
  geom_line(mapping = aes(y = h_3(t), colour = "h = t2-1"), linewidth = 1.25) +
  scale_colour_manual("Funciones",values=c("black","red"))
```



El caso simple también significa que podemos reescribir nuestra definición original como

$$DF = \sum_i \partial_i f(W(h_1), \dots, W(h_n)) D(W(h_i))$$

que se ve como una regla de la cadena normal. No voy a demostrarlo, pero la regla del producto también es exactamente como lo que esperamos:

$$D(FG) = FD(G) + GD(F)$$

Ahora, usaremos otro ejemplo simple, un movimiento browniano al cuadrado:

$$F(t) = (W_t)^2$$

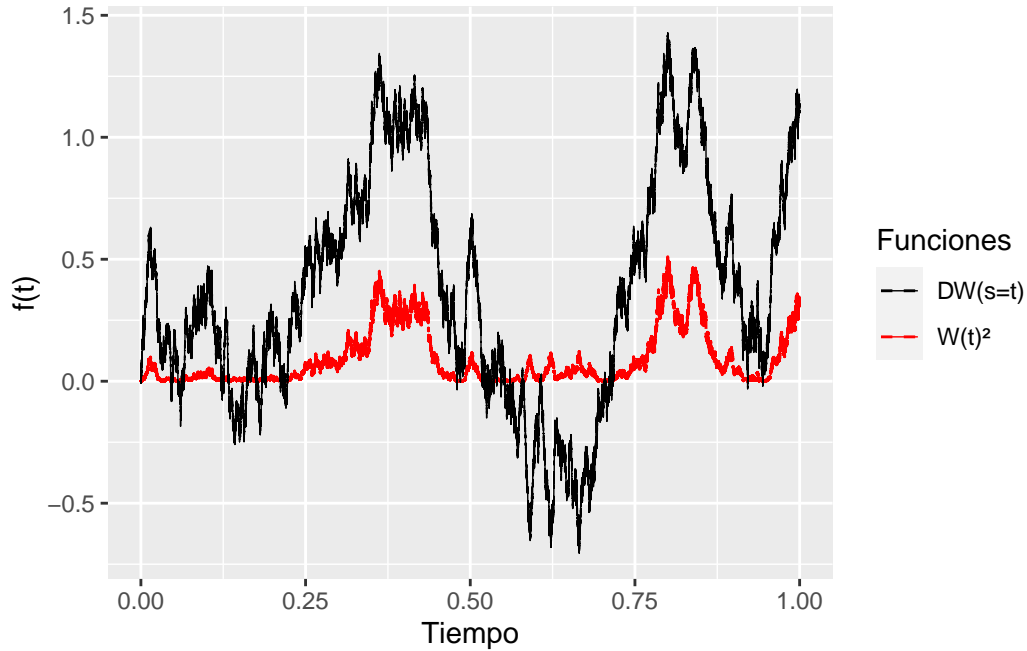
En este caso, usamos $h(s) = 1_{[0,t]}(s)$. Observemos que en realidad hay dos variables relacionadas con el tiempo: una es la t original a la que se refiere $F(t)$, y la segunda es la introducida por la derivada de Malliavin, a la que llamaremos s . Con esto en mente, estamos listos para brindar una respuesta:

$$D_s F = 2 W_t (DW_t) = 2 W_t h = 2 W_t 1_{[0,t]}(s)$$

Hay dos cosas para destacar acá. En primer lugar, la derivada toma valores distintos de cero solo cuando $0 < s < t$, y esto tiene sentido: no esperamos que la función F cambie más allá del tiempo t en el que se evalúa. En segundo lugar, esto también significa que cuando graficamos $W(h(t))$ y $h(t)$ juntos, fue un poco impreciso porque estábamos asumiendo $s = t$ para poder dibujarlos juntos. Aun así, creo que ese supuesto nos permite ver que hay cierta conexión entre la derivada de Malliavin y la “varianza instantánea” o volatilidad del proceso: cuanto más cerca está la derivada de 0, las perturbaciones o movimientos son menores en el proceso estocástico. Graficaré ahora una trayectoria de ejemplo para F y DF , con $t \in [0, 1]$ y $s = t$ (que, nuevamente, no es la imagen completa, pero me parece útil para conectar los puntos):

```
# Setup
steps = 100000
h <- 1/steps
t <- seq(from = 0, to = 1, length.out=steps)
dWt <- c(0, rnorm(n=(steps-1),mean = 0,sd = sqrt(h)))
Wt <- cumsum(dWt)
f <- Wt**2
Df <- 2*Wt

# Plot
ggplot(mapping = aes(x = t)) +
  xlab('Tiempo') +
  ylab('f(t)') +
  geom_line(mapping = aes(y = f, colour="W(t)2"),linetype = 2) +
  geom_line(mapping = aes(y = Df, colour="DW(s=t)"), linewidth = 0.25) +
  scale_colour_manual("Funciones",values=c("black","red"))
```

Integral de Skorohod y más integración por partes

Dijimos que los operadores lineales vienen con un operador adjunto. Las derivadas de Malliavin no son la excepción y, dado que definimos D con un producto interno y una función h en $\langle DF, h \rangle_H$, buscaremos un operador adjunto, al que llamamos δ , de modo que:

$$\langle DF, h \rangle_H = \langle F, \delta(h) \rangle_{R^n}$$

Podemos encontrarlo haciendo un poco de manipulación, tomando algunos atajos y renunciando a cierto rigor matemático. A partir de la definición anterior y suponiendo que $h = h_1$, tomamos el valor esperado del producto interno, realizamos integración por partes y algunos reemplazos y obtendremos:

$$\begin{aligned}
\mathbb{E} [\langle DF, h \rangle_H] &= \mathbb{E} \left[\sum_i \partial_i f \langle h_i, h \rangle_H \right] \\
&= \int_{\mathbb{R}^n} \underbrace{\partial_1 f(x)}_{dv} (2\pi)^{-n/2} \underbrace{e^{-\|x\|^2/2}}_u dx \\
&= - \int_{\mathbb{R}^n} f(x) (-x_1) \underbrace{(2\pi)^{-n/2} e^{-\|x\|^2/2} dx}_{\text{Medida gaussiana}} \\
&= \int_{\mathbb{R}^n} x_1 f(x) d\mu_1 \\
&= \mathbb{E} [F W(h_1)] \\
&= \mathbb{E} [F W(h)] \\
&= \mathbb{E} \left[F \int_0^1 h dB \right]
\end{aligned}$$

Esto significa que la esperanza matemática de la variable aleatoria DF medida usando la función h entre 0 y 1 como regla, tiene el mismo valor que la esperanza de F multiplicado por esa integral de h a la derecha. Estamos dando a entender que el operador δ es esa integral. A esta integral se la llama la **integral de Skorohod**, y coincide con una integral de Ito para procesos no anticipativos³:

$$\delta(h) = \int_0^1 h dB_t$$

A estas alturas, probablemente hayas notado lo común que es invertir integrales y derivadas mediante integración por partes.

¿Y ahora qué?

Por fin llegamos a una definición de la derivada de Malliavin. Nos ha llevado un tiempo, pero ahora estamos en una posición cómoda para usarla para algo útil.

³Formalmente, se trata de procesos predecibles o adaptados. Un proceso predecible tiene un valor en t conocido en momentos anteriores a t (por ejemplo, una función no aleatoria). Un proceso adaptado tiene un valor en t conocido en t o posterior (por ejemplo, el precio de una acción). Los procesos no adaptados, o procesos anticipativos, solo se conocen después de que haya pasado algún tiempo, y su información depende del futuro (por ejemplo, que una granizada se produzca en los próximos 7 días).

Alos, & Lorite, E. 2021. *Malliavin Calculus in Finance: Theory and Practice (1st ed.)*. 1.^a ed. Financial Mathematics Series. Chapman; Hall/CRC. <https://doi.org/10.1201/9781003018681>.

Friz, Peter K. 2002. «An Introduction to Malliavin Calculus». En. <https://api.semanticscholar.org/CorpusID:2479628>.