Caracterización geométrica de la independencia en un espacio gaussiano

POR DAVID NUALART (*)

Conferencia impartida en la Academia
el día 4 de marzo de 1992

1. Introducción

La noción de independencia juega un papel fundamental en la teoría de la probabilidad. Aunque a primera vista aparece como un concepto intuitivo y elemental, puede dar lugar a problemas bastante complejos. Nos proponemos en esta conferencia de poner en relación la independencia, que es una noción básicamente probabilística, con la noción de ortogonalidad, que es de carácter geométrico. Recordaremos primero algunas nociones básicas.

En un espacio de probabilidad (Ω, F, P), dos variables aleatorias reales $F, G: \Omega \to R$ se dice que son *independientes* si se cumple

$$P(a \leq F \leq b, c \leq G \leq d) = P(a \leq F \leq b) P(c \leq G \leq d),$$

para todo par de intervalos $[a, b]$ y $[c, d]$. Esto es equivalente a decir que la distribución de probabilidad que el vector aleatorio (F, G) induce en el plano R^2 es igual al producto de las distribuciones marginales de las variables F y G. La independencia de las variables F y G puede caracterizarse de diversas maneras que proporcionan criterios útiles para decidir si F y G son independientes. Por ejemplo, es bien sabido que F y G son independientes si y sólo si la densidad $f_{F,G}(x, y)$ del vector aleatorio (F, G), en caso de que exista, se escribe como producto de las densidades marginales. También la independencia de F y G equivale a que la función característica del vector (F, G) se exprese como producto de las funciones características de F y G.

En el caso de dos variables F y G de segundo orden (es decir, de cuadrado integrable), tales que $E(F) = E(G) = 0$, la independencia de F y G implica su ortogonalidad en el espacio de Hilbert $L^2(\Omega, F, P)$. En

(*) Facultad de Matemáticas. Universidad de Barcelona

efecto, la independencia implica que la esperanza de F por G es igual al producto de las esperanzas de F y G, y por tanto es igual a cero. La ortogonalidad es, por tanto, una propiedad más débil que la independencia. Por otra parte, si el vector aleatorio (F, G) tiene una ley gaussiana centrada, entonces, la condición $E(FG) = 0$ implica la independencia entre F y G. Esto es debido a que la densidad conjunta de F y G se puede expresar de la forma

$$f_{F,G}(x, y) = \frac{1}{2\pi \det \Gamma} \exp\left(-\frac{1}{2}(x, y)\, \Gamma^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \right)$$

donde $\Gamma = \begin{pmatrix} E(F^2) & E(FG) \\ E(FG) & E(G^2) \end{pmatrix}$ es la matriz de covarianzas del vector (F, G). Es claro que la densidad $f_{F,G}(x, y)$ factoriza en el producto de una función de x por una función de y si Γ es diagonal lo que equivale a $E(FG) = 0$.

2. Independencia en un espacio gaussiano de dimensión finita

Supongamos que el espacio de probabilidad es $\Omega = R^n$ y la probabilidad P es la ley normal $N(0, I_n)$. Ello quiere decir que las coordenadas, que designaremos por $x_1, \ldots, x_n$, son variables aleatorias independientes y con distribución $N(0, 1)$. En este contexto, dos variables aleatorias reales serán dos funciones medibles

$$F = F(x_1, \ldots, x_n), \quad G = G(x_1, \ldots, x_n).$$

Si las funciones F y G son lineales, entonces la distribución conjunta de F y G es gaussiana y la independencia entre F y G equivale a la ortogonalidad de los gradientes de las funciones F y G:

$$F \text{ es independiente de } G \iff E(FG) = 0 \iff \nabla F \cdot \nabla G = 0$$

En efecto, si $F(x_1, \ldots, x_n) = \sum_{i=1}^n a_i x_i$ y $G(x_1, \ldots, x_n) = \sum_{i=1}^n b_i x_i$, se tiene

$$E(FG) = \sum_{i=1}^{n} a_i b_i = \nabla F \cdot \nabla G .$$

Nos podemos preguntar entonces la siguiente cuestión. Dadas dos funciones F y G no necesariamente lineales, pero continuamente diferenciables en un abierto V de probabilidad uno, cual es la relación entre las dos propiedades siguientes:

[A] F y G son independientes

[B] $(\nabla F \cdot \nabla G)(x) = 0$, para todo $x \in V$.

La propiedad [B] significa que las hipersuperficies F = cst. y G = cst. son ortogonales. Existen ejemplos simples en los que [A] y [B] se verifican. Por ejemplo si hacemos un cambio de coordenadas polares a las variables x_1 y x_2, las funciones

$$F = arctan \frac{x_1}{x_2}, \quad G = \sqrt{x_1^2 + x_2^2}$$

son independientes y sus gradientes son ortogonales. Sin embargo, en general las propiedades [A] y [B] no son comparables. Como ilustración mencionaremos los ejemplos siguientes:

Ejemplo 1:

Tomemos $F = \max(x_1, x_2)$ y $G = \min(x_1, x_2)$. En este caso las variables F y G no son independientes ya que $F \geq G$, pero sus gradientes son ortogonales en el abierto $\{x_1 \neq x_2\}$ ya que

$$\nabla F = (1_{\{x_1 > x_2\}}, 1_{\{x_1 < x_2\}}, 0, \ldots, 0)$$

$$\nabla G = (1_{\{x_1 < x_2\}}, 1_{\{x_1 > x_2\}}, 0, \ldots, 0)$$

Ejemplo 2:

Definamos en primer lugar (suponiendo $n \geq 4$)

$$X = \frac{1}{2\pi} arctan \frac{x_1}{x_2}, \quad Y = \frac{1}{2\pi} arctan \frac{x_3}{x_4}.$$

Las variables X y Y son independientes y con distribución uniforme en el intervalo [0, 1]. A partir de X y de Y construimos las variables F y G de la manera siguiente

$$F = \sqrt{-\log X} \cos 2\pi Y, \quad G = \sqrt{-\log X} \, sen \, 2\pi Y.$$

Las variables F y G son independientes y con distribución gaussiana $N(0, 1)$ pero $\nabla F \cdot \nabla G \neq 0$ casi por todo. Observemos que la transformación $(X, Y) \to F$ es la utilizada usualmente por las calculadoras para simular variables con ley gaussiana $N(0, 1)$ a partir de números aleatorios, es decir de variables con distribución uniforme en [0, 1].

Recientemente, Ustunel y Zakai [5] han resuelto el problema de la caracterización de la independencia en el caso de funciones polinómicas. Para explicar el resultado fundamental de estos autores es preciso introducir algunas notaciones. Designaremos por H_1 el subespacio de

$L^2(R^n, N(0,I_n))$ formado por las funciones lineales. Para cada $m \geq 1$ representaremos por H_m el espacio de los polinomios de grado m que son ortogonales en $L^2(R^n, N(0,I_n))$ a todo polinomio de grado $m-1$. Finalmente H_0 será el espacio de las constantes. Se tiene la descomposición en suma ortogonal

$$L^2(R^n, N(0,I_n)) = \bigoplus_{m=0}^{\infty} H_m \tag{2.1}$$

El resultado de Ustunel y Zakai dice lo siguiente

Teorema 1. *Consideremos dos polinomios $F \in H_p$, $G \in H_q$. Entonces las dos propiedades son equivalentes:*

[A] *F y G son independientes.*

[C] $\sum_{k=1}^n (\nabla_{i_1} \cdots \nabla_{i_{p-1}} \nabla_k F)(\nabla_{j_1} \cdots \nabla_{j_{q-1}} \nabla_k G) = 0$, *para todo conjunto de índices $\{i_1, \ldots, i_{p-1}, j_1, \ldots, j_{q-1}\} \in \{1,\ldots,n\}$, donde ∇_i representa la derivada parcial respecto la variable x_i.*

Observación : La condición [C] implica la propiedad [B] de ortogonalidad de los gradientes de F y G. Más precisamente, para dos polinomios $F \in H_p$, $G \in H_q$, la ortogonalidad de los gradientes equivale a la anulación de la simetrización en las variables $(i_1, \ldots, i_{p-1}, j_1, \ldots, j_{q-1})$ del tensor que aparece en la condición [C]. En consecuencia, para dos polinomios de este tipo la independencia implica la ortogonalidad de los gradientes, y se obtiene una respuesta parcialmente afirmativa a la cuestión que nos preguntábamos antes.

Demostración del teorema 1:
Sabemos que una base ortonormal del espacio H_m viene dada por los polinomios de Hermite multidimensionales definidos como

$$H_\alpha(x) = \prod_{i=1}^n H_{i(\alpha)}(x_i),$$

donde $\alpha = \{\alpha_1, \ldots, \alpha_m\} \subset \{1, \ldots, n\}$, i (α) representa el número de coordenadas de α iguales a i, y $H_n(x)$ es el n-ésimo polinomio de Hermite. Entonces todo polinomio $R \in H_m$ se expresa de la forma

$$R = \sum_\alpha A_\alpha H_\alpha,$$

donde $A_\alpha = E(RH_\alpha) = \sqrt{m!}\nabla_{\alpha_1} \dots \nabla_{\alpha_m} R$. Con estos elementos indicaremos seguidamente como se demuestra el teorema.

(i) Es claro que, a partir de la caracterización de la independencia de funciones lineales mediante la ortogonalidad de sus gradientes, la condición [C] equivale a la independencia de las σ-álgebras τ_F y τ_G generadas respectivamente por las funciones lineales

$$\{\nabla_{i_1} \dots \nabla_{i_{p-1}} F, \quad 1 \leq i_1, \dots, i_{p-1} \leq n\}, \tag{2.2}$$

y

$$\{\nabla_{j_1} \dots \nabla_{j_{q-1}} G, \quad 1 \leq j_1, \dots, j_{q-1} \leq n\}. \tag{2.3}$$

Entonces la implicación [C] $\Rightarrow$ [A] se deduce de que las variables F y G son τ_F y τ_G medibles, respectivamente. Esto último puede verse de la forma siguiente. Sea R_F el subespacio de R^n generado por los gradientes de las funciones lineales que aparecen en (2.2). Podemos hacer un cambio de base ortonormal en R^n completando una base de R_F. Supongamos que las nuevas coordenadas son $y_1, \dots, y_n$, donde $y_1, \dots, y_k$ son funciones lineales de las x_i que pertenecen a R_F. Entonces se cumple que

$$\frac{\partial}{\partial y_r} \nabla_{i_1} \dots \nabla_{i_{p-1}} F = \langle \nabla y_r , \nabla_{i_1} \dots \nabla_{i_{p-1}} \rangle = 0,$$

para todo $k < r \leq n$, debido al carácter ortogonal de la nueva base. Esto nos dice que en la representación de F como combinación lineal de polinomios de Hermite multidimensionales en las variables y_j, solamente aparecerán las coordenadas y_j con índices $1 \leq j \leq k$, y por lo tanto, F es τ_F medible.

(ii) Recíprocamente, supongamos que F y G son independientes. Para simplificar la demostración supondremos que $p = q = 2$. Entonces, F y G serán de la forma

$$F = \sum_{i \leq j} a_{ij} (x_i x_j - \delta_{ij}), \quad G = \sum_{i \leq j} b_{ij} (x_i x_j - \delta_{ij}).$$

La condición [C] que queremos probar nos dice en este caso particular que

$$\sum_{k=1}^{n} a_{i_1 k} b_{j_1 k} = 0$$

para todo par de índices i_1, j_1. Para comprobar esta condición utilizaremos la igualdad

$$E(F^2 G^2) = E(F^2)\, E(G^2), \tag{2.4}$$

que se deduce de la independencia entre F y G. Por una parte es fácil ver que

$$E(F^2)\,E(G^2) = 4\sum_{i\leq j} a_{ij}^2 \sum_{i\leq j} b_{ij}^2 .$$

Por otra parte, mediante un cálculo un poco más complicado se llega a que

$$E(F^2G^2) = 4\sum_{i\leq j} a_{ij}^2 \sum_{i\leq j} b_{ij}^2 + c\sum_{i,j=1}^{n}\left|\sum_{k=1}^{n} a_{ik}b_{jk}\right|^2 + d \ ,$$

donde $c > 0$ y $d \geq 0$ son dos constantes, y de ello se deduce que el coeficiente de c debe ser cero. En el caso general la demostración se basa también en la igualdad (2.4) pero los cálculos son más difíciles.

El análisis de la relación entre la independencia y la ortogonalidad de los gradientes se ha proseguido en dos direcciones:

(i) Independencia de dos sigma-álgebras σ y τ (véase [4,6]).
(ii) Independencia condicional [3].

Describiremos seguidamente los resultados obtenidos en la dirección (i). La noción de independencia se generaliza, como es bien sabido, al caso de dos subsigma-álgebras σ y τ contenidas en la sigma-álgebra F. Se dice que σ y τ son independientes si

$$P(A \cap B) = P(A)\,P(B),$$

para todo par de sucesos $A \in \sigma$, $B \in \tau$. Se plantea entonces el problema de relacionar la independencia de las sigma-álgebras con propiedades geométricas como la ortogonalidad. Para ello es preciso introducir la noción de *espacio tangente*.

Consideremos una sigma-álgebra τ generada con una colección numerable $\{g_i\,,\, i \geq 1\}$ de funciones que son infinitamente diferenciables en un abierto $V \subset R^n$ de probabilidad uno. Una tal sigma-álgebra la denominaremos *regular*. El *espacio tangente* a una sigma-álgebra regular τ será por definición el espacio vectorial

$$K_\tau(x) = \langle \nabla g_i(x),\, i \geq 1 \rangle, \quad x \in V.$$

En realidad esta notación no es muy apropiada porque si τ está generada por una única variable g entonces el espacio tangente a las hipersuperficies de nivel de g es precisamente el ortogonal de $K_\tau(x)$. Sin embargo, ésta es la definición que se maneja en la literatura sobre este tema y la que nosotros utilizaremos. En principio el espacio tangente puede depender del sistema de generadores. Existen condiciones suficientes para que el espacio tangente sea universal. Como ejemplo de tales condiciones podemos citar las siguientes (véase [4]):

(i) El operador de proyección del gradiente de una función en el ortogonal de K_τ es un operador cerrado en $L^2(R^n, P)$.

(ii) El conjunto $\{\nabla g_i(x), i \geq 1\}$ es denso en $L^2(\tau)$ (conjunto de variables de cuadrado integrable y τ-medibles) respecto la norma del espacio de Sobolev $W^{1,2}$ (espacio de las variables de cuadrado integrable tales que sus gradientes son también de cuadrado integrable respecto P).

La noción de espacio tangente permite traducir propiedades probabilísticas en propiedades analíticas. Por ejemplo, si X es una variable τ medible que pertenece al espacio de Sobolev $W^{1,2}$, y suponemos que se cumple la condición (i), entonces ∇X pertenece al espacio tangente casi seguramente.

En el caso particular de dos sigma-álgebras τ_1 y τ_2 regulares generadas por funciones *lineales*, se tiene que

(2.5) τ_1 y τ_2 son independientes $\Leftrightarrow$ $K_{\tau_1} \perp K_{\tau_2}$.

Este tipo de sigma-álgebras puede caracterizarse de la forma siguiente:

Proposición 1. *Una sigma-álgebra regular $\tau = \sigma\{g_i, i \geq 1\}$ está generada por funciones lineales si y sólo si su espacio tangente es constante casi seguramente y para todo vector v del espacio tangente la función lineal $\langle v, x \rangle$ es τ-medible.*

En esta caracterización el hecho de que el espacio tangente sea constante no es suficiente para que la sigma-álgebra esté generada por funciones lineales. Por ejemplo la sigma-álgebra generada por x_1^2 tiene un espacio tangente constante pero no se puede generar mediante funciones lineales.

Como en el caso de las variables aleatorias nos podemos preguntar si la equivalencia anterior se generaliza a sigma-álgebras regulares generadas por funciones no lineales. Sin embargo, la construcción de sigma-álgebras independientes que tengan generadores no lineales es un problema difícil, y solamente en determinadas situaciones particulares se pueden dar respuestas. En los trabajos [3,6] se han introducido hipótesis de estabilidad para poder generalizar la equivalencia (2.5). Se dice que una sigma-álgebra τ_1 es estable si el espacio $L^2(\tau_1)$ es invariante por el operador de proyección en H_n, para todo $n \geq 1$. Por ejemplo la sigma-álgebra generada por x_1^2 es estable. Citaremos seguidamente dos de los resultados básicos obtenidos en las referencias [3,6].

Proposición 2. *Consideremos dos sigma-álgebras* τ_1 y τ_2 *regulares, tales que* τ_1 *es estable. Entonces,*

$$K_{\tau_1} \perp K_{\tau_2}\,, \text{ c.s.} \quad \Rightarrow \quad \tau_1 \text{ y } \tau_2 \text{ son independientes.}$$

Proposición 3. *Consideremos tres sigma-álgebras* τ_1, τ_2 y τ *regulares tales que* τ *tiene generadores lineales, y* τ_1 *es* τ*-estable (es decir,* τ_1 *es estable respecto a la ley condicionada por* τ*). Entonces*

$K_{\tau_1} \ominus K_\tau \perp K_{\tau_2} \ominus K_\tau$, c.s. $\Rightarrow$ τ_1 y τ_2 son condicionalmente independientes dada τ.

3. Cálculo diferencial en espacios gaussianos de dimensión infinita

En la situación descrita en el apartado anterior la probabilidad gaussiana estaba definida en R^n, que es un espacio de dimensión finita en el que tenemos a nuestra disposición el cálculo diferencial clásico. En general nos interesará considerar una medida gaussiana definida en un espacio de dimensión infinita, como un espacio de Banach. De forma quizás más abstracta se puede definir un espacio gaussiano como un espacio de probabilidad (Ω, F, P) en el que existe una colección numerable y gaussiana (las leyes conjuntas son normales) de variables aleatorias reales $\{X_i, i \geq 1\}$ que genera la sigma-álgebra F. Aunque a priori Ω no posea ninguna estructura de variedad diferenciable, es posible desarrollar un cálculo diferencial en el espacio (Ω, F, P), siguiendo las ideas introducidas por Malliavin, Segal, Ito, entre otros. Un cálculo diferencial de este tipo será útil para tratar diversos problemas y en particular, nos permitirá generalizar la relación entre ortogonalidad e independencia presentada en el apartado anterior.

El ejemplo más interesante de espacio gaussiano de dimensión infinita es el espacio de Wiener. Vamos a restringir nuestro análisis a este ejemplo. Tomaremos por lo tanto, $\Omega = C_0([0, 1])$, y la probabilidad P será la ley del movimiento browniano. Es decir, P es una probabilidad en la sigma-álgebra de Borel de Ω tal que el proceso estocástico canónico $W_t(\omega) = \omega(t)$, $0 \leq t \leq 1$, es un proceso gaussiano de media cero y covarianza $E(W_t W_s) = s \wedge t$.

Una variable aleatoria $F : \Omega \to R$ diremos que es elemental si es de la forma

$$F(\omega) = f(\omega(t_1), \ldots, \omega(t_m)), \quad f \in C_b^\infty(R^m).$$

Se define el gradiente de F como el proceso estocástico $\nabla F = \{\nabla_t F, t \in [0, 1]\}$ dado por

$$\nabla_t F = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} (\omega(t_1), \ldots, \omega(t_m))\, \mathbf{1}_{[0,t_i]}(t).$$

Es decir, en comparación con el gradiente finito-dimensional introducido en el apartado anterior, las coordenadas son ahora los puntos de intervalo [0,1], y el espacio R^n se sustituye por el espacio de Hilbert $L^2([0,1])$. La relación entre este gradiente y la derivada de Fréchet de F viene dada por la ecuación siguiente:

$$\int_0^1 \nabla_t F\, \dot{h}_t\, dt = \frac{d}{d\varepsilon} F(\omega + \varepsilon h)|_{\varepsilon=0}, \quad \dot{h} \in L^2([0,1])$$

En 1976 Malliavin utilizó el cálculo diferencial en el espacio de Wiener con objeto de dar una demostración probabilística al teorema de hipoelipticidad de Hörmander. El cálculo de Malliavin se ha revelado como una herramienta potente para tratar diversos problemas del análisis estocástico. El objetivo inicial de Malliavin fue la obtención de criterios para que una variable aleatoria $F: \Omega \to R^d$ tuviese densidad diferenciable respecto a la medida de Lebesgue. Para ello, partiendo de la definición anterior de gradiente, se introdujeron los espacios de Sobolev en Ω, de forma parecida a como se hace en el análisis clásico en dimensión finita. Es decir, para todo $k \geq 0$ y todo real $p > 1$, se define el espacio $D^{p,k}$ como la adherencia del conjunto de variables elementales respecto a la seminorma

$$\| F \|_{p,k} = \| F \|_p + \sum_{j=1}^{k} \| \nabla^j F \|_{L^p(\Omega; L^2([0,1]^j))}$$

donde ∇^j representa la aplicación iterada del gradiente.

Los espacios Sobolev en el espacio Wiener han sido estudiados por Watanabe en [7]. Es interesante ver cómo se pueden deducir propiedades probabilísticas de variables aleatorias en el espacio de Wiener a partir de propiedades analíticas. Consideremos, por ejemplo, un vector aleatorio $F = (F^1, \ldots, F^d)$. Entonces se verifican las siguientes propiedades.

(I) Si las componentes F^i pertenecen al espacio $D^{1,2}$ para todo $i = 1, \ldots, d$, entonces, el soporte de la distribución de probabilidad de F en R^d es un conjunto cerrado y conexo. En particular, esto implica que la ley de F no puede ser discreta. Si además $d = 1$ y $E(|F|^{-4}) < \infty$ entonces, F es una variable positiva o negativa, es decir $P(F \geq 0)$ vale cero o uno.

(II) Se define la *matriz de Malliavin* del vector F, suponiendo que las componentes de F pertenecen al espacio $D^{1,2}$, como

$$(\sigma_F)_{ij} = \langle DF^i, DF^j \rangle_{L^2([0,1])}, \quad 1 \leq i, j \leq d.$$

Entonces si $\det \sigma_F > 0$ c.s., la distribución de probabilidad de F es absolutamente continua.

(III) Si $F = (F^1, \ldots, F^d)$ es un vector aleatorio cuyas componentes están en todos los espacios $D^{p,k}$ y el inverso del determinante de la matriz de Malliavin tiene momentos de todos los órdenes, entonces, F tiene una densidad infinitamente diferenciable.

Cuando el vector aleatorio F es el valor de un proceso de difusión X_t en un instante $t > 0$, el criterio (III) permite demostrar el teorema de Hörmander que asegura que la ley de X_t tiene una densidad infinitamente diferenciable si los coeficientes de la ecuación satisfecha por X son C^∞ y cumplen una condición de no degeneración (condición de Hörmander). Más precisamente consideremos un proceso de movimiento browniano en R^k. $W = \{(W_t^1, \ldots, W_t^k), t \geq 0\}$, así como una familia de funciones $A_i : R^d \to R^d$, $0 \leq i \leq k$, que supondremos infinitamente diferenciable con todas sus derivadas acotadas. Entonces, fijada una condición inicial $x_0 \in R^d$, podemos definir el proceso de difusión d-dimensional solución de la ecuación diferencial estocástica siguiente:

$$dX_t = \sum_{i=1}^{k} A_i(X_t)\, dW_t^i + A_0(X_t)\, dt, \quad X_0 = x_0. \tag{3.1}$$

Las funciones A_i dan lugar a los operadores diferenciales $\sum_{j=1}^{d} A_i^j(x) \frac{\partial}{\partial x_j}$. El cálculo estocástico de Ito nos asegura que la ley de probabilidad p_t del vector aleatorio $X(t)$, $t > 0$ satisface el problema de Cauchy siguiente

$$\frac{\partial p_t}{\partial t} = L^*(p_t), \quad t > 0, \quad p_0 = \delta_{x_0}, \tag{3.2}$$

donde L es el operador diferencial de segundo orden (y L^* designa su adjunto), definido por

$$L = \frac{1}{2} \sum_{i=1}^{k} (A_i)^2 + A_0.$$

Se observa a partir de la ecuación (3.2), denominada ecuación de Kolmogorov, que la hipoelipticidad del operador L está relacionada con la propiedad de que p_t tenga densidad infinitamente diferencible. En este sentido, el teorema de Hörmander, que asegura que el operador L es hipoelíptico, tiene una traducción probabilística en el hecho que la ley de probabilidad de una difusión tenga densidad infinitamente diferenciable. Consideremos la función $\tilde{A}_0 = A_0 - \frac{1}{2}\sum_{i=1}^{k} A_i^{\nabla} A_i$, donde $A^{\nabla} B$ representa la derivada covariante de B en la dirección de A. Entonces, como aplicación del criterio general (III), se puede probar el siguiente resultado:

Teorema 1. *Supongamos que los coeficientes de la ecuación (3.1) cumplen la siguiente condición de Hörmander:*
(H): El álgebra de Lie engendrada por los operadores diferenciales A_i , $[A_i , \tilde{A}_0]$, $1 \leq i \leq k$ *(donde $[\cdot]$ representa el paréntesis de Lie), tiene dimensión máxima igual a d en el punto x_0, condición inicial de la ecuación (3.1).*
Entonces la ley del vector aleatorio $X(t)$ tiene una densidad infinitamente diferenciable, para todo instante $t > 0$.

4. Independencia en el espacio de Wiener

Utilizando el cálculo de Malliavin, que nos proporciona una noción de gradiente, se pueden generalizar los resultados obtenidos en el apartado 2 al caso de variables aleatorias y sigma-álgebras en el espacio de Wiener. Los elementos del espacio H_p, en lugar de ser combinaciones lineales de polinomios de Hermite, serán ahora integrales estocásticas múltiples:

$$F = I_p(f) = p!\int_{\{t_1 < \ldots < t_p\}} f(t_1,\ldots,t_p)\, dW(t_1)\ldots dW(t_p) ,$$

donde f es una función simétrica del espacio $L^2([0,1])$. Estas integrales estocásticas múltiples fueron introducidas por Ito y verifican la propiedad siguiente de isometría:

$$E(I_p(f)^2) = p!\|f\|^2 .$$

La descomposición ortogonal (2.1) es ahora la denominada descomposición del espacio $L^2(C_o([0,1]), P)$ en *Caos de Wiener.*

Consideremos dos integrales estocásticas múltiples $I_p(f)$ y $I_q(g)$. Los resultados del apartado 2 se traducen en las siguientes propiedades:

(i) $I_p(f)$ y $I_q(g)$ son independientes $\Leftrightarrow f\otimes_1 g = 0$, casi por todo.
(ii) $\langle\nabla(I_p(f)), \nabla(I_q(g))\rangle_{L^2([0,1])} = 0 \Leftrightarrow f\tilde{\otimes}_1 g = 0$, casi por todo.

En las anteriores propiedades, $f\otimes_1 g = 0$ denota la contracción de un índice entre las funciones f y g definida como la función de $L^2([0,1]^{p+q-2})$ dada por la relación

$$f\otimes_1 g = \int_0^1 f(\cdot,t)\, g(\cdot,t)\, dt ,$$

y $f\tilde{\otimes}_1 g$ es la simetrización de $f\otimes_1 g$.

Como ilustración de la simplicidad de cálculos cuando se trabajo con integrales múltiples demostraremos la equivalencia (ii).

Demostración de (ii). En primer lugar debemos indicar que el operador gradiente opera sobre las integrales múltiples de una forma muy simple. En efecto, el proceso estocástico $\nabla_t(I_p(f))$ es igual a $p\, I_{p-1}(f(\cdot,t))$. Por lo tanto el producto escalar de gradientes que aparecen en la propiedad (ii) se escribirá como

(4.1)

$$\langle \nabla(I_p(f)), \nabla(I_q(g))\rangle_{L^2([0,1])} = pq\int_0^1 I_{p-1}(f(\cdot,t))\, I_{q-1}(f(\cdot,t))\, dt$$

Por otra parte el producto de dos integrales múltiples se puede expresar como combinación lineal de integrales múltiples de órdenes inferiores de la manera siguiente:

(4.2)

$$I_{p-1}(f(\cdot,t))I_{q-1}(f(\cdot,t)) = \sum_{r=0}^{(p-1)\wedge(q-1)} r!\binom{p-1}{r}\binom{q-1}{r} I_{p+q-2r-2}(f(\cdot,t)\tilde{\otimes}_r g(\cdot,t))$$

donde $\tilde{\otimes}_r$ representa la simetrización de la contracción de r índices. Entonces, utilizando la fórmula (4.2) vemos que la condición $\langle \nabla(I_p(f)), \nabla(I_q(g))\rangle_{L^2([0,1])} = 0$ equivale a que la suma de integrales múltiples

$$\sum_{r=0}^{(p-1)\wedge(q-1)} r!\binom{p-1}{r}\binom{q-1}{r} I_{p+q-2r-2}(f\tilde{\otimes}_{r+1} g)$$

se anule. Como las integrales estocásticas de órdenes diferentes son ortogonales entre sí, que la suma anterior se anule equivale a que cada término sea 0 y teniendo en cuenta la definición de las contracciones, esto quiere decir que $f\tilde{\otimes}_1 g$ debe ser cero.

En cuanto a la propiedad (i) la implicación $\Leftarrow$ se demuestra como en el caso de dimensión finita, mientras que la implicación contraria se deduce, como en el caso de dimensión finita, de la igualdad

$E(I_p(f)^2 I_q(g)^2) = E(I_p(f)^2)\, E(I_q(g)^2)$, y de la utilización de la fórmula del producto de dos integrales múltiples.

Análogamente, la noción de espacio tangente se puede introducir en el espacio de Wiener y se puede estudiar la relación entre la independencia de sigma-álgebras y la ortogonalidad de sus espacios tangentes. Pueden establecerse resultados análogos a las Proposiciones 2 y 3 del apartado 2; introduciendo la noción de sigma-álgebra estable. La noción de estabilidad introducida en el apartado 2 puede caracterizarse mediante la estabilidad respecto a cierto operador en el espacio $L^2(\Omega)$. Consideremos el operador L definido a partir de la descomposición en caos de Wiener, de la forma siguiente:

$$L\left(\sum_{p=0}^{\infty} I_p(f_p)\right) = \sum_{p=1}^{\infty} p\, I_p(f_p).$$

El operador L es un operador autoadjunto y no acotado en $L^2(\Omega)$, que genera el denominado semigrupo de Ornstein-Uhlenbeck. Existe la siguiente relación entre el operador L y el gradiente introducido anteriormente:

$$(4.3) \qquad E(\langle \nabla F, \nabla G \rangle_{L^2([0,1])}) = E(LF\, G) = E(LG\, F),$$

si F y G son dos variablers aleatorias en el espacio de Sobolev $D^{2,2}$. El operador L tiene un inverso, que designaremos por L^{-1}, que actua sobre las variables centradas, mediante la multiplicación por el factor $1/p$ en cada caos de Wiener de orden p. Se trata de un operador acotado. Entonces, se demuestra que una sigma-álgebra τ es estable si y solo si el espacio $L_0^2(\tau)$ (espacio de las variables aleatorias centradas, de cuadrado integrable y τ medibles) es invariante respecto al operador L^{-1}

Veamos, para finalizar, como se demuestra la proposición 2 del apartado 2 mediante las técnicas del cálculo de Malliavin. Para demostrar la independencia de las dos sigma-álgebras τ_1 y τ_2 es suficiente con ver que dadas dos variables aleatorias centradas F y G, τ_1 y τ_2 medibles, respectivamente, y del espacio D^∞ (lo que nos permitirá aplicar sin problemas los operadores ∇ y L), se cumple

$$(4.4) \qquad E(FG) = 0$$

Para demostrar la igualdad (4.4), escribimos $F = L(L^{-1}F)$ (lo cual es posible porque F es centrada). Utilizando (4.3) tendremos

$$(4.5) \qquad E(FG) = E[L(L^{-1}F).G] = E[\langle \nabla(L^{-1}F), \nabla(G)\rangle].$$

Debido a la estabilidad de la sigma-álgebra τ_1, la variable aleatoria $L^{-1}F$ es τ_1 medible. De ello se deduce que su gradiente pertenece al espacio tangente K_{τ_1}, y entonces, la ortogonalidad de los espacios tangentes implica que el producto escalar que aparece en (4.5) debe ser cero.

Un estudio detallado de la sigma-álgebras que poseen determinadas propiedades de estabilidad ha sido realizado en la referencia [4].

Bibliografía

[1] O. KALLENBERG: On an independence criterion for multiple Wiener integrals. *Annals of Probability*, **19**, 483-485 (1991).

[2] P. MALLIAVIN: Stochastic calculus of variation and hypoelliptic operators. Proc. Int. Symp. on S.D.E., 1976, Kyoto, ed. K. Ito, North Holland.

[3] D. NUALART, A.S. USTUNEL: Geometric analysis of conditional independence on Wiener space. *Probab. Theory Rel. Fields*, **89**, 407-422 (1991).

[4] D. NUALART, A.S. USTUNEL and M. ZAKAI: Some relations among classes of σ-fields on Wiener space. *Probab. Theory and Rel. Fields*, **85**, 119-129 (1990).

[5] A.S. USTUNEL and M. ZAKAI: On the independence and conditioning on Wiener space. *Annals of Probability*, **17**, 1441-1453 (1989).

[6] A.S. USTUNEL and M. ZAKAI: On the structure of independence on Wiener space. *Journal of Functional Analysis*, **90**, 113-137 (1990).

[7] S. WATANABE: Stochastic differential equations and Malliavin calculus. Tata Institute of Fundamental Research. Springer-Verlag, 1984.