

Universidad EAFIT
Departamento de Ciencias Básicas

Ecuaciones Diferenciales Estocásticas y casos de aplicación en Finanzas

Freddy H. Marín Sánchez

Trabajo de Investigación presentado como requisito parcial para optar
por el título de Magister en Matemáticas Aplicadas

Director
Ulises Cárcamo

Departamento de Ciencias Básicas
UNIVERSIDAD EAFIT
Medellín

Agradecimientos

El desarrollo de este trabajo fue posible gracias al invaluable apoyo del departamento de Ciencias Básicas de la Universidad EAFIT, al programa de Ingeniería Matemática y a todo el esfuerzo de los profesores de la maestría en Matemáticas Aplicadas.

Quiero agradecer también a todas aquellas personas que de una u otra forma hicieron posible la realización de este trabajo, a mi familia y en especial a los profesores Ulises Cárcamo, Hermilson Velásquez, José Valdés y Jairo Villegas; así como a mis compañeros Francisco Zuluaga, Fredy Pérez y al estudiante Tomás Olarte.

Índice general

Agradecimientos	I
Prefacio	IX
1. Conceptos Preliminares	1
1.1. Introducción	1
1.2. Espacio de Probabilidad	2
1.3. Esperanza Condicional	4
1.3.1. Propiedades elementales de la esperanza condicional	4
1.4. Procesos Estocásticos	5
2. Movimiento Browniano e Integrales Estocásticas	7
2.1. Introducción	7
2.2. Integrales Estocásticas	10
2.2.1. La Fórmula de Itô	13
3. Ecuaciones Diferenciales Estocásticas	15
3.1. Existencia y Unicidad de las Soluciones	16
3.2. Ecuaciones Diferenciales Estocásticas Lineales	17
3.2.1. Fórmula Estocástica de Liouville	18
3.2.2. Fórmula de Variación de las Constantes	19
3.2.3. Ecuaciones Lineales Escalares	20
3.2.4. Ecuaciones Lineales en el Sentido Estricto	21
3.2.5. Ecuaciones Lineales Autónomas	21

3.3.	Aproximaciones Numéricas para EDE	22
3.3.1.	Método de Aproximación de Euler	22
3.3.2.	Método de Aproximación de Milstein	25
3.4.	Ecuaciones Diferenciales Parciales	26
3.4.1.	El Problema de Cauchy	27
3.4.2.	Las Ecuaciones de Kolmogorov	31
4.	Dinámica de Portafolios: una aplicación financiera	35
4.1.	Portafolios	35
4.1.1.	Dividendos	39
4.2.	Precios en condiciones de arbitraje	41
4.2.1.	Reclamos contingentes y arbitraje	42
4.2.2.	Justificación Semi-intuitiva	44
5.	La Fórmula de Black-Scholes: otra aplicación financiera	47
5.1.	Ecuación de Black-Scholes	47
5.2.	Valoración de riesgo neutral	51
5.3.	La fórmula de Black-Scholes	53
6.	Procesos AR(1), GARCH(1,1) y de Reversión a la Media	57
6.1.	Modelos con Heterocedasticidad Condicional	57
6.1.1.	El proceso AR(1)	58
6.1.2.	El proceso GARCH(1,1)	59
6.2.	Procesos de Reversión a la Media	60
6.2.1.	Procesos de Reversión a la Media de un solo factor	60
7.	Estimación de Modelos de Difusión: la conexión con el mundo “real”	63
7.1.	Método de Máxima Verosimilitud	63
7.1.1.	Estimación de Máxima Verosimilitud para el Proceso de Ornstein- Uhlenbeck	64

7.1.2.	Estimadores de Máxima Verosimilitud	64
7.2.	Método de Momentos y Generador Infinitesimal	66
7.2.1.	Condiciones de Momento	66
7.2.2.	Identificación	67
7.2.3.	Método de Momentos	68
7.2.4.	Descomposición Espectral del Generador Infinitesimal	69
7.2.5.	Descomposición Canónica no Lineal	70
7.2.6.	Estimación para la Descomposición Espectral	71
7.2.7.	Estimación basada en el Kernel	72
7.3.	Métodos basados en las funciones de escala y la velocidad	72
7.3.1.	Escala intrínseca y tiempo discreto	72
7.3.2.	Estimación de observaciones de valor discreto	74
7.4.	Método de Momentos Simulados	75
7.5.	Inferencia Indirecta	77
7.6.	Método de los Momentos Eficientes	78
7.6.1.	Aplicación a modelos de tiempo continuo	78
8.	Aplicación: el caso del aluminio	81
8.1.	Estimadores de Máxima Verosimilitud	83
8.2.	Simulación	85
8.3.	El caso del Aluminio	87
8.4.	Conclusiones	90

Índice de figuras

2.1. Una trayectoria de un Movimiento Browniano Estándar	8
2.2. Izquierda: función diferenciable. Derecha: Función no diferenciable en $x = 0$, existen infinitos tangentes que se pueden cruzar por el punto.	8
2.3. Movimientos Brownianos en los cuales se resalta la forma similar a escala.	9
2.4. Trayectoria de un Movimiento Browniano Geométrico.	10
3.1. Aproximaciones de Euler para $dX_t = \lambda X_t dt + \mu X_t dB_t$ con $\lambda = 1, \mu = 2$	24
3.2. Aproximaciones de Milstein para $dX_t = \lambda X_t dt + \mu X_t dB_t$ con $\lambda = 1, \mu = 2$	26
4.1. Con $K = 100$, función de contrato call Europea (a la izquierda) y función de contrato put Europea (a la derecha).	43
5.1. Black-Scholes para una opción call Europea	56
8.1. Parámetros de cada simulación	86
8.2. Simulaciones para cuatro procesos GARCH	86
8.3. Gráfica de la serie de aluminio	87
8.4. Autocorrelación del aluminio	87
8.5. Gráfica de los residuales para la serie del aluminio	88
8.6. Modelo GARCH para la serie del aluminio	88
8.7. Gráfica del modelo GARCH	89

Prefacio

El cálculo estocástico tiene muchas aplicaciones en las áreas de Biología, Sociología, Demografía, Economía y Finanzas. Por otro lado, las series de tiempo son una herramienta estadística muy utilizada en el estudio de series financieras y económicas. Estas dos herramientas pueden emplearse conjuntamente para estudiar la dinámica de precios de algunos commodities con características especiales. En este trabajo se presentan de manera formal los elementos matemáticos de las Ecuaciones Diferenciales Estocásticas con un enfoque dirigido a la Dinámica de Portafolios y la Valoración de Opciones según Black-Scholes. Debido a que algunos precios de commodities, como el aluminio, presentan efectos GARCH, se propone hacer una fusión entre procesos de reversión a la media del tipo Ornstein Uhlenbeck y los procesos GARCH.

En el primer capítulo se presentan algunos preliminares matemáticos que servirán como insumo teórico para el desarrollo de los demás capítulos.

En el capítulo 2, se hará una descripción formal del Movimiento Browniano Estándar y de las Integrales Estocásticas con sus propiedades fundamentales y, por último, se estudiara la fórmula de Itô multi-dimensional.

Para el siguiente capítulo se establecen los criterios de existencia y unicidad de las soluciones de EDE y se presentarán, de manera formal, las ecuaciones de Kolmogorov para EDP, la representación estocástica de Feynman-Kac y las soluciones explícitas de EDE matriciales con las aproximaciones numéricas de Euler y Milstein para el caso unidimensional.

Los capítulos 4 y 5, Dinámica de Portafolios y La Fórmula de Black-Scholes respectivamente, hacen referencia al contexto financiero en el cual pueden ser aplicados los conceptos estudiados en los capítulos anteriores.

Los capítulos 7 y 8 son claves para la implementación de los procesos de reversión a la media con efecto GARCH, en los que se utiliza el Método de Máxima Verosimilitud para la estimación de los parámetros del proceso de reversión a la media del tipo Ornstein Uhlenbeck y del proceso GARCH(1,1). También se hace una breve revisión de otros métodos de estimación de modelos de difusión importantes, aunque son un poco más complejos al momento de su implementación.

En el último capítulo se lleva a cabo una aplicación sencilla basada en resultados previos de

simulaciones realizadas para algunas series con características apropiadas, donde se utiliza la combinación de series de tiempo no lineales y las Ecuaciones Diferenciales Estocásticas para el estudio de la dinámica de precios diarios del aluminio.

Capítulo 1

Conceptos Preliminares

1.1. Introducción

En este capítulo se hará una breve revisión de las notaciones y elementos básicos de la teoría de la probabilidad, la teoría de la medida y de los procesos estocásticos.

También se darán algunas definiciones, teoremas y proposiciones importantes que servirán como base teórica para la presentación de definiciones y teoremas en los capítulos posteriores.

Definición 1.1.1. Un experimento se dice que es *aleatorio* si su resultado no puede ser determinado de antemano. Para desarrollar un modelo matemático adecuado, los posibles resultados de un experimento aleatorio deben estar bien definidos y ser observables.

Definición 1.1.2. El conjunto Ω de todos los resultados de un experimento aleatorio se llama *espacio muestral* y los elementos $\omega \in \Omega$ se llaman *puntos muestrales*.

Definición 1.1.3. Sea Ω un conjunto arbitrario. Una familia \mathfrak{F} de subconjuntos de Ω es una σ -álgebra sobre Ω si se cumplen las siguientes condiciones:

- I) $\emptyset \in \mathfrak{F}$ y $\Omega \in \mathfrak{F}$ donde \emptyset representa el conjunto vacío
- II) $\forall A \in \mathfrak{F} \Rightarrow A^c \in \mathfrak{F}$ donde A^c es el complemento de A en Ω
- III) $A_1, A_2, \dots \in \mathfrak{F} \Rightarrow A := \bigcup_{i=1}^{\infty} A_i \in \mathfrak{F}$
- IV) $A_1, A_2, \dots \in \mathfrak{F} \Rightarrow A := \bigcap_{i=1}^{\infty} A_i \in \mathfrak{F}$

Nota 1. De este modo una σ -álgebra sobre Ω es una familia de subconjuntos de Ω que contiene a Ω y es cerrado bajo complementación bajo la formación de uniones contables y bajo la formación de intersecciones contables. Además, la cerradura bajo complementación

y bajo la formación de uniones contables implican cerradura bajo la formación de intersecciones contables puesto que $\bigcap A_i = (\bigcup A_i^c)^c$. De esta forma se puede definir una σ -álgebra sólo con las condiciones **i**), **ii**) y **iii**).

Definición 1.1.4. Si \mathcal{U}_1 y \mathcal{U}_2 son dos σ -álgebras sobre Ω tales que $\mathcal{U}_1 \subseteq \mathcal{U}_2$, se dice que \mathcal{U}_1 es una sub σ -álgebra de \mathcal{U}_2 .

Definición 1.1.5. Sea \mathcal{C} una familia de subconjuntos de Ω , existe una σ -álgebra minimal $\sigma(\mathcal{C})$ sobre Ω que contiene a \mathcal{C} . A $\sigma(\mathcal{C})$ se le llama σ -álgebra generada por \mathcal{C} .

Definición 1.1.6. Si $\Omega = \mathbb{R}^n$ y \mathcal{C} es la familia de todos los conjuntos abiertos en \mathbb{R}^n , entonces $\mathfrak{B}^n = \sigma(\mathcal{C})$ se conoce como σ -álgebra de Borel y los elementos de \mathfrak{B}^n se llaman *conjuntos de Borel*.

En general una función que toma valores reales $X : \Omega \rightarrow \mathbb{R}$ se dice que es \mathfrak{F} -medible si $\{\omega : X(\omega) \leq a\} \in \mathfrak{F}$ para toda $a \in \mathbb{R}$.

La función X también se conoce con el nombre de variable aleatoria de valor real (\mathfrak{F} -medible). Una función de valor en \mathbb{R}^n $X(\omega) = (X_1(\omega), X_2(\omega), \dots, X_n(\omega))^T$ se dice \mathfrak{F} -medible si todos los elementos X_i son \mathfrak{F} -medibles.

Lema 1.1.1. Si $X, Y : \Omega \rightarrow \mathbb{R}^n$ son dos funciones dadas, entonces Y es $\sigma(X)$ -medible si y sólo si existe una función Borel-medible $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $Y = g(X)$.

1.2. Espacio de Probabilidad

Definición 1.2.1. Una medida de probabilidad sobre un espacio medible (Ω, \mathfrak{F}) es una función $p : \mathfrak{F} \rightarrow [0, 1]$ tal que:

I) $p(\emptyset) = 0$

II) $p(\Omega) = 1$

III) Si $A_1, A_2, \dots \in \mathfrak{F}$ y $\{A_i\}_{i=1}^{\infty}$ es una sucesión de conjuntos disjuntos dos a dos entonces,

$$p\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} p(A_i)$$

Definición 1.2.2. La tripleta $(\Omega, \mathfrak{F}, p)$ se llama *espacio de probabilidad*. Si $(\Omega, \mathfrak{F}, p)$ es un espacio de probabilidad, tómesese $\tilde{\mathfrak{F}} = \{A \subset \Omega : \exists B, C \in \mathfrak{F} \text{ tales que } B \subset A \subset C, p(B) = p(C)\}$. Entonces $\tilde{\mathfrak{F}}$ es una σ -álgebra y se conoce como la completación de \mathfrak{F} .

Definición 1.2.3. Si $\mathfrak{F} = \tilde{\mathfrak{F}}$, se dice que el espacio de probabilidad $(\Omega, \mathfrak{F}, p)$ es completo, si no, se puede extender p a $\tilde{\mathfrak{F}}$ definiendo $p(A) = p(B) = p(C)$ para $A \in \tilde{\mathfrak{F}}$, donde $B, C \in \mathfrak{F}$ con las propiedades $B \subset A \subset C$ y $p(B) = p(C)$. $(\Omega, \tilde{\mathfrak{F}}, p)$ es un espacio de probabilidad completo llamado completación de $(\Omega, \mathfrak{F}, p)$.

¹La T simboliza la transpuesta de la matriz

Definición 1.2.4. Si X es una variable aleatoria de valor real y es integrable con respecto a la medida de probabilidad p , el número $E[X] = \int_{\Omega} X(\omega) dp(\omega)$ se llama *esperanza* de X con respecto a p . Adicionalmente el número $V[X] = E(X - E[X])^2$ se llama *varianza* de X .

Si Y es otra variable aleatoria de valor real, $\text{cov}[X, Y] = E[(X - E[X])(Y - E[Y])]$ se llama *covarianza* de X y Y .

Si $\text{cov}[X, Y] = 0$ se dice que X y Y están incorrelacionados. Para una variable aleatoria de valores en \mathbb{R}^n , $X = (X_1, X_2, \dots, X_n)^T$ se define $E[X] = (E[X_1], E[X_2], \dots, E[X_n])^T$, análogamente para una variable aleatoria matricial $X = (X_{ij})_{n \times m}$ de dimensión $n \times m$ se define $E[X] = (E[X_{ij}])_{n \times m}$.

Si X y Y son ambas variables aleatorias de valores en \mathbb{R}^n , la matriz simétrica definida no negativa de dimensión $n \times n$ dada por $\text{cov}[X, Y] = E[(X - E[X])(Y - E[Y])^T]$ se llama *matriz de covarianzas*.

Cuando se relacionan mediante la covarianza variables aleatorias cuyos valores son de escalas muy diferentes, es recomendable estandarizar la covarianza con el fin de lograr una relación que se pueda interpretar de manera más adecuada. Luego se tiene el concepto de *correlación* entre dos variables aleatorias X_1 y X_2 , el cual se define como

$$\text{corr}[X_1, X_2] := \frac{\text{cov}[X_1, X_2]}{\sigma_{X_1}\sigma_{X_2}} = \frac{E[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})]}{\sigma_{X_1}\sigma_{X_2}}.$$

Definición 1.2.5. El número $E[X]^p$ con $p > 0$ se llama el p -ésimo momento de X . Para $p \in (0, \infty)$, sea $L^p = L^p(\Omega, \mathbb{R}^n)$ la familia de variables aleatorias X con valores en \mathbb{R}^n con $E[X]^p < \infty$.

A continuación se presentaran cuatro conceptos de convergencia muy importantes.

Sean X y X_k , con $k \geq 1$, variables aleatorias con valores en \mathbb{R}^n .

- I) Si existe un conjunto P -nulo (Conjunto con medida de probabilidad $P = 0$), $\Omega_0 \in \mathfrak{F}$ tal que para cada $\omega \notin \Omega_0$, la sucesión $\{X_k(\omega)\}$ converge a $X(\omega)$ en el sentido usual en \mathbb{R}^n , entonces $\{X_k\}$ se dice que converge a X casi seguramente (c.s.) o con probabilidad 1, y se escribe $\lim_{k \rightarrow \infty} X_k = X$ c.s.
- II) Si para cada $\epsilon > 0$, $P\{\omega : |X_k(\omega) - X(\omega)| > \epsilon\} \rightarrow 0$ conforme $k \rightarrow \infty$, entonces se dice que $\{X_k\}$ converge a X estocásticamente o en probabilidad.
- III) Si X_k y X pertenecen a L^p y $E|X_k - X|^p \rightarrow 0$, entonces se dice que $\{X_k\}$ converge a X en el p -ésimo momento o en L^p .
- IV) Si para cada función g acotada y continua definida sobre \mathbb{R}^n , $\lim_{k \rightarrow \infty} E_g(X_k) = E_g(X)$, entonces se dice que $\{X_k\}$ converge a X en distribución.

Teorema 1.2.1 (Teorema de convergencia Monótona). *Si $\{X_k\}$ es una sucesión decreciente de variables aleatorias no negativas, entonces*

$$\lim_{k \rightarrow \infty} E[X_k] = E(\lim_{k \rightarrow \infty} X_k)$$

Teorema 1.2.2 (Teorema de convergencia Dominada²). *Sea $p \geq 1$, $\{X_k\} \subset L^p(\Omega; \mathbb{R}^n)$ y $Y \in L^p(\Omega; \mathbb{R})$. Suponga que $|X_k| \leq Y$ c.s. y que $\{X_k\}$ converge a X en probabilidad. Entonces $X \in L^p(\Omega; \mathbb{R}^n)$*

Definición 1.2.6. Dos conjuntos $A, B \in \mathfrak{F}$ se dice que son *independientes* si $P(A \cap B) = P(A)P(B)$. Tres conjuntos $A, B, C \in \mathfrak{F}$ se dice que son *independientes* si $P(A \cap B) = P(A)P(B)$; $P(A \cap C) = P(A)P(C)$; $P(B \cap C) = P(B)P(C)$ y $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Definición 1.2.7. Si X y Y son dos variables aleatorias independientes e integrables de valor real, entonces XY también es integrable y además $E[XY] = E[X]E[Y]$.

Si $X, Y \in L^2(\Omega; \mathbb{R})$ están incorrelacionados entonces $V[X + Y] = V[X] + V[Y]$.

Definición 1.2.8. Sean $A, B \in \mathfrak{F}$ con $P(B) > 0$. La *probabilidad condicional* de A condicionado a B está dada por $P(A/B) = \frac{P(A \cap B)}{P(B)}$.

Proposición 1.2.3. *Si X_1 y X_2 son dos variables aleatorias independientes, entonces $\text{cov}[X_1, X_2] = 0$.*

1.3. Esperanza Condicional

Definición 1.3.1. Sea $X \in L^1(\Omega; \mathbb{R})$. Sea $\mathcal{G} \subset \mathfrak{F}$ una sub σ -álgebra de \mathcal{F} , de este modo el espacio $(\Omega; \mathcal{G})$ define un espacio medible. En general X no es necesariamente \mathcal{G} -medible. Existe una variable aleatoria Y \mathcal{G} -medible única c.s. tal que $Y = E[X/\mathcal{G}]$ y se llama *esperanza condicional* de X bajo la condición \mathcal{G} .

Si \mathcal{G} es la σ -álgebra generada por la variable aleatoria Y , se puede escribir $E[X/\mathcal{G}] = E[X/Y]$.

1.3.1. Propiedades elementales de la esperanza condicional

A continuación se presentarán algunas propiedades relevantes de la esperanza condicional.

Sea Ω un conjunto arbitrario y \mathcal{G} una σ -álgebra sobre Ω .

a) $E[E[X/\mathcal{G}]] = E[X]$

²Cuando Y es acotado, este teorema también se conoce como el teorema de convergencia acotado.

- b) $|E[X/\mathcal{G}]| \leq E[|X|/\mathcal{G}]$ c.s.
- c) Si $\mathcal{G} = \{\Phi, \Omega\}$ entonces $E[X/\mathcal{G}] = E[X]$
- d) Si $X \geq 0$ entonces $E[X/\mathcal{G}] \geq E[X]$
- e) Si X es \mathcal{G} -medible entonces $E[X/\mathcal{G}] = X$
- f) Si $X = C = \text{constante}$ entonces $E[X/\mathcal{G}] = C$
- g) Sean $a, b \in \mathbb{R}$, $E[aX + bY/\mathcal{G}] = aE[X/\mathcal{G}] + bE[Y/\mathcal{G}]$
- h) Si $X \leq Y$ entonces $E[X/\mathcal{G}] \leq E[Y/\mathcal{G}]$
- i) Si X es \mathcal{G} -medible entonces $E[XY/\mathcal{G}] = X \cdot E[Y/\mathcal{G}]$, en particular, $E[E[X/\mathcal{G}]Y/\mathcal{G}] = E[X/\mathcal{G}]E[Y/\mathcal{G}]$
- j) Si $\sigma(X), \mathcal{G}$ son independientes entonces $E[X/\mathcal{G}] = E[X]$
- k) Si $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathfrak{F}$ entonces $E[E[X/\mathcal{G}_2]/\mathcal{G}_1] = E[X/\mathcal{G}_1]$
- l) Sea $X = (X_1, X_2, \dots, X_n)^T \in L^1(\Omega; \mathbb{R}^n)$, su esperanza condicional bajo \mathcal{G} está definida por $E[X/\mathcal{G}] = (E[X_1/\mathcal{G}], \dots, E[X_n/\mathcal{G}])^T$.

1.4. Procesos Estocásticos

Definición 1.4.1. Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad. Una *filtración* es una familia $\{\mathfrak{F}_t\}_{t \geq 0}$ creciente de sub σ -álgebras de \mathfrak{F} , es decir, $\mathfrak{F}_s \subset \mathfrak{F}_t \subset \mathfrak{F}$ para toda $0 \leq s \leq t \leq \infty$.

La filtración se dice continua a derecha si $\mathfrak{F}_S = \bigcap_{t > S} \mathfrak{F}_t$ para toda $S \geq 0$.

Cuando el espacio de probabilidad es completo, se dice que la filtración satisface las condiciones usuales siempre que sea continua a derecha y \mathfrak{F}_0 contenga todos los conjuntos P -nulos.

Definición 1.4.2. Una familia $\{X_t\}_{t \in I}$ de variables aleatorias de valores en \mathbb{R}^n se llama *proceso estocástico* con parámetro t y espacio de estado \mathbb{R}^n .

Habitualmente el parámetro $t \in I$ es la semirrecta $\mathbb{R}_+ = [0, \infty)$, pero en algunas ocasiones también puede ser un intervalo $[a, b]$ de enteros no negativos.

Definición 1.4.3. Se dice que un proceso estocástico $\{X_t\}$ es $\{\mathfrak{F}_t\}$ -adaptado o simplemente *adaptado* si para cada t , X_t es \mathfrak{F}_t -medible.

Definición 1.4.4 (Martingala). Un proceso $\{\mathcal{M}_t\}_{t=0}$ integrable \mathfrak{F} -adaptado de valores en \mathbb{R}^n se dice que es una *martingala* con respecto a $\{\mathfrak{F}_t\}$ (o simplemente martingala), si $E[\mathcal{M}_t/\mathfrak{F}_s] = \mathcal{M}_s$ para toda $s, t \in [0, \infty)$.

Definición 1.4.5 (Vector Aleatorio Gaussiano). Un *vector aleatorio normal* o Gaussiano es un vector que sigue una distribución normal o Gaussiana.

La distribución normal n -dimensional o Gaussiana es dada por la densidad:

$$f_X(\mathbf{X}) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{X} - \mu)\Sigma^{-1}(\mathbf{X} - \mu)'\right\}; \mathbf{X} \in \mathbb{R}^n$$

con parámetros $\mu \in \mathbb{R}^n$ y Σ , donde $\mu = \mu_X$ es la esperanza de \mathbf{X} .

Además Σ es una matriz $n \times n$ simétrica definida positiva, Σ^{-1} es la inversa y $\det(\Sigma)$ es el determinante de Σ .

Como la densidad de un vector aleatorio Gaussiano multivariado \mathbf{X} está dada por el parámetro μ que es la esperanza μ_X de \mathbf{X} y Σ es la matriz de covarianza Σ_X , la densidad de un vector Gaussiano está completamente determinada por su esperanza y su matriz de covarianza. En particular, si $\mu = \vec{0}$ y Σ es la matriz identidad n -dimensional I_n , tenemos que $\det(I_n) = 1$ y $\Sigma^{-1} = I_n$. La densidad f_X es simplemente el producto de n densidades normales estándar: $f_X = (x_1, x_2, \dots, x_n) = \varphi(x_1)\varphi(x_2) \dots \varphi(x_n)$.

Se escribe $N(\mu, \Sigma)$ para la distribución de un vector aleatorio Gaussiano n -dimensional \mathbf{X} con esperanza μ y matriz de covarianza Σ .

Definición 1.4.6. Las distribuciones de dimensión finita (fidis) de un proceso estocástico \mathbf{X} son todas las distribuciones de los vectores de dimensión finita $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ con $n = 1, 2, \dots$ y $t_1, t_2, \dots, t_n \in \mathbb{T}$.

Definición 1.4.7 (Proceso Gaussiano). Un proceso estocástico se llama *proceso Gaussiano* si todas fidis son Gaussianas multivariadas. Por tanto la distribución de un proceso estocástico Gaussiano está determinado solamente por la colección de las esperanzas y las matrices de covarianza de las fidis.

Un proceso Gaussiano *simple* sobre $\mathbb{T} = [0, 1]$ consiste de variables aleatorias independientes e idénticamente distribuidas (iid) con una distribución Normal con media 0 y varianza 1 $N(0, 1)$. En este caso las fidis están caracterizadas por las funciones de distribución

$$\begin{aligned} P(X_{t_1} \leq x_1, X_{t_2} \leq x_2, \dots, X_{t_n} \leq x_n) &= P(X_{t_1} \leq x_1) P(X_{t_2} \leq x_2) \dots P(X_{t_n} \leq x_n) \\ &= \Phi(x_1)\Phi(x_2) \dots \Phi(x_n). \end{aligned}$$

Capítulo 2

Movimiento Browniano e Integrales Estocásticas

2.1. Introducción

En este capítulo se abordarán algunos conceptos centrales que son muy importantes en la teoría de los procesos estocásticos, la física y las finanzas.

Primero se hará una descripción formal del Movimiento Browniano, junto con sus principales características, además de algunos procesos asociados al Movimiento Browniano. Luego se definirá la integral estocástica con sus propiedades, y por último se presentará formalmente un resultado de suma importancia para el cálculo estocástico y sus aplicaciones, el lema de Itô.

Definición 2.1.1. Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad con una filtración $\{\mathfrak{F}_t\}_{t \geq 0}$. Un movimiento Browniano Estándar unidimensional $\{B_t\}_{t \geq 0}$, es un proceso \mathfrak{F}_t -adaptado, continuo de valor real con las siguientes propiedades:

- I) $B_0 = 0$.
- II) El proceso $\{B_t\}_{t \geq 0}$ tiene incrementos *estacionarios*, es decir, $B_t - B_s \stackrel{d}{=} B_{t+h} - B_{s+h}$ ¹ para toda $s, t \in T$ y para toda $s + h, t + h \in T$
- III) El proceso $\{B_t\}_{t \geq 0}$ tiene incrementos independientes, es decir, para cada elección de $t_i \in T$ con $t_1 < t_2 < \dots < t_n$ y $n \geq 1$, $B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ son variables aleatorias independientes.
- IV) Para $0 \leq s < t < \infty$, el incremento $B_t - B_s$ sigue una distribución normal con media cero y varianza $t - s$.

¹La notación $B_t - B_s \stackrel{d}{=} B_{t+h} - B_{s+h}$ establece una identidad distribucional entre los dos términos, en general $B_t - B_s \neq B_{t-s}$

v) $\{B_t\}_{t \geq 0}$ tiene trayectorias continuas y es en ningún punto diferenciable.

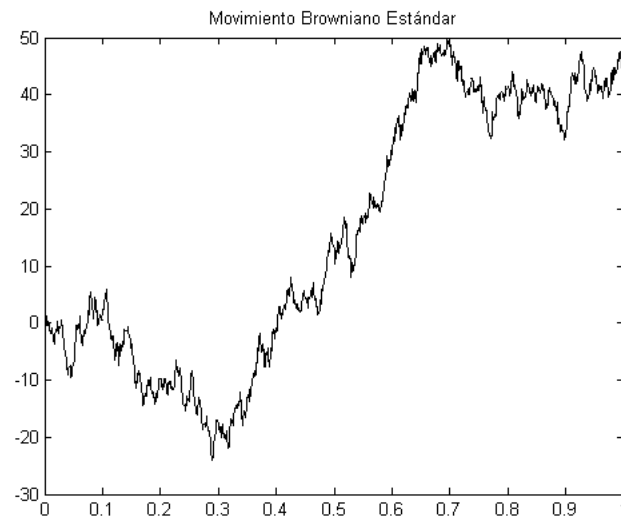


Figura 2.1: Una trayectoria de un Movimiento Browniano Estándar

En una función diferenciable, cada punto x de la gráfica puede ser aproximado por una función lineal única que es tangente a la función en dicho punto x . Para el caso del Movimiento Browniano, existen infinitas rectas tangentes a la “curva” en cada punto, hecho que compromete la diferenciable de la trayectoria browniana.

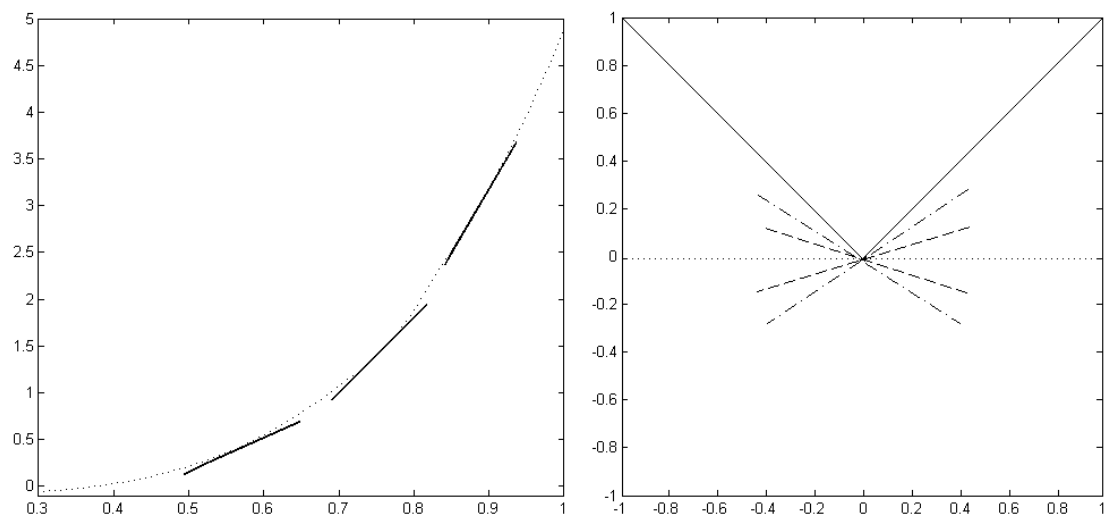


Figura 2.2: Izquierda: función diferenciable. Derecha: Función no diferenciable en $x = 0$, existen infinitas tangentes que se pueden cruzar por el punto.

Algunas propiedades relevantes del Movimiento Browniano:

- 1) El Movimiento Browniano es un proceso Gaussiano.

II) La covarianza del Movimiento Browniano es $\text{cov}(B_s, B_t) = \min\{s, t\}$.

III) El Movimiento Browniano es 0.5-autosimilar, es decir,

$$(T^{\frac{1}{2}}B_{t_1}, \dots, T^{\frac{1}{2}}B_{t_n}) \stackrel{d}{=} (B_{Tt_1}, \dots, B_{Tt_n}).$$

Cuando se habla de *autosimilaridad*, se quiere decir que los patrones a escala de una trayectoria browniana en algún intervalo de tiempo grande o pequeño, conserva una forma similar, pero no son idénticos. En la 2.3 se muestra gráficamente el concepto intuitivo de ser autosimilar.

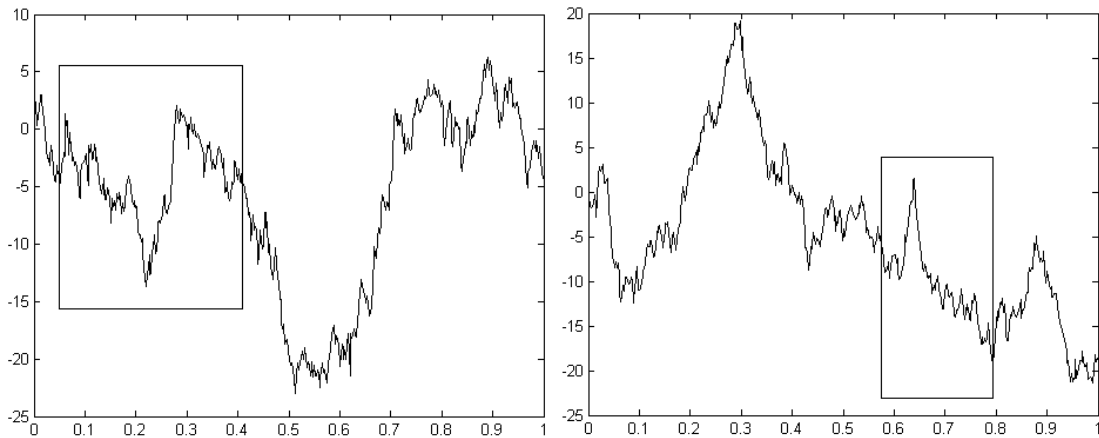


Figura 2.3: Movimientos Brownianos en los cuales se resalta la forma similar a escala.

Existen algunos procesos asociados al Movimiento Browniano, como por ejemplo el Movimiento Browniano Bridge, el Movimiento Browniano Drift, y uno conocido por su relación con la dinámica de precios de un activo, el Movimiento Browniano Geométrico.

Definición 2.1.2. El proceso $\{X_t\}_{t \geq 0}$ dado por $X_t = \exp(\alpha t + \lambda B_t)$ donde B_t es un Movimiento Browniano Estándar, $\lambda > 0$ y $\alpha \in \mathcal{R}$; se conoce como *Movimiento Browniano Geométrico*.

El Movimiento Browniano Geométrico no es un proceso Gaussiano y tiene las siguientes características:

- I) $E[X_t] = \exp(\alpha + \frac{\lambda^2}{2})t$
- II) $V[X_t] = \exp(2\alpha + \lambda^2)t(\exp(\lambda^2 t) - 1)$
- III) $\text{cov}(X_t, X_s) = \exp(\alpha + \frac{\lambda^2}{2})(t + s)(\exp \lambda^2 s - 1)$.

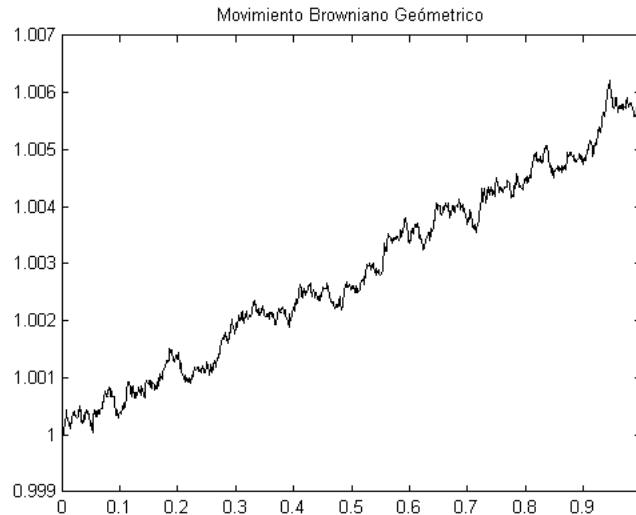


Figura 2.4: Trayectoria de un Movimiento Browniano Geométrico.

2.2. Integrales Estocásticas

En esta sección se definirá la integral estocástica $\int_0^t f(s)dB_s$ con respecto al Movimiento Browniano B_t . Como se ha dicho anteriormente, una trayectoria browniana es en ningún punto diferenciable y la integral no puede ser definida en la forma habitual. Sin embargo, se puede definir la integral para una extensa clase de procesos estocásticos usando la naturaleza estocástica del Movimiento Browniano. Esta integral fue definida por Kiyoshi Itô en 1949, y se conoce como Integral de Estocástica de Itô.

Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad con una filtración $\{\mathfrak{F}_t\}_{t \geq 0}$.

Sea $\{B_t\}_{t \geq 0}$ un Movimiento Browniano unidimensional definido sobre el espacio de probabilidad adaptado a la filtración.

Definición 2.2.1. Sea $0 \leq a < b < \infty$. Se denota por $\mathcal{M}^2([a, b]; \mathbb{R})$ al espacio de todos los procesos $f = \{f(t)\}_{a \leq t \leq b}$ medibles y $\{\mathfrak{F}\}$ -adaptados de valor real tales que $\|f\|_{a,b}^2 := E \left[\int_a^b |f(t)|^2 dt \right] < \infty$ ².

Definición 2.2.2. Un proceso estocástico de valor real $g = \{g(t)\}_{a \leq t \leq b}$ se llama *proceso simple*, si existe una partición $a = t_0 < t_1 < \dots < t_k = b$ en $[a, b]$, y variables aleatorias acotadas $\xi_i, 0 \leq i \leq k - 1$ tal que ξ_i es \mathfrak{F}_{t_i} -medible y

$$g(t) = \xi_0 I_{[t_0, t_1]}(t) + \sum_{i=1}^{k-1} \xi_i I_{(t_i, t_{i+1}]}(t) \quad (2.1)$$

Se denota por $\mathcal{M}_0([a, b]; \mathbb{R})$ a la familia de todos estos procesos, donde $\mathcal{M}_0([a, b]; \mathbb{R}) \subset \mathcal{M}^2([a, b]; \mathbb{R})$

² $\|\cdot\|_{a,b}$ define una métrica sobre $\mathcal{M}^2([a, b]; \mathbb{R})$ y el espacio es completo bajo esta métrica.

Definición 2.2.3 (Integral de Itô parte I). Para un proceso simple g de la forma 2.1 en $\mathcal{M}_0([a, b]; \mathbb{R})$, se define la integral de Itô como

$$\int_a^b g(t)dB_t = \sum_{i=0}^{k-1} \xi_i (B_{t_{i+1}} - B_{t_i})$$

Definición 2.2.4 (Media Cero e Isometría de Itô para procesos simples). Sea $g \in \mathcal{M}_0([a, b]; \mathbb{R})$ entonces se tiene que

- I) $E[\int_a^b g(t)dB_t] = 0$
 II) $E|\int_a^b g(t)dB_t|^2 = E[\int_a^b |g(t)|^2 dt]$

Demostración. 2.2.4

- I) Como ξ_i es \mathfrak{F}_{t_i} -medible y además $B_{t_{i+1}} - B_{t_i}$ es independiente de \mathfrak{F}_{t_i} , $E[\int_a^b g(t)dB_t]$ es igual a $\sum_{i=0}^{k-1} E[\xi_i (B_{t_{i+1}} - B_{t_i})] = \sum_{i=0}^{k-1} E[\xi_i]E[(B_{t_{i+1}} - B_{t_i})] = 0$.
- II) Nótese que $B_{t_{j+1}} - B_{t_j}$ es independiente de $\xi_i \xi_j (B_{t_{i+1}} - B_{t_i})$ si $i < j$. De este modo

$$\begin{aligned} E \left| \int_a^b g(t)dB_t \right|^2 &= \sum_{0 \leq i, j \leq k-1} E[\xi_i \xi_j (B_{t_{i+1}} - B_{t_i}) (B_{t_{j+1}} - B_{t_j})] \\ \sum_{i=0}^{k-1} E[\xi_i^2 (B_{t_{i+1}} - B_{t_i})^2] &= \sum_{i=0}^{k-1} E[\xi_i^2] E[(B_{t_{i+1}} - B_{t_i})^2] \\ &= \sum_{i=0}^{k-1} E[\xi_i^2] E[(t_{i+1} - t_i)^2] \\ &= E \int_a^b |g(t)|^2 dt \end{aligned}$$

■

Lema 2.2.1 (Linealidad de la Integral de Itô). Sean $g_1, g_2 \in \mathcal{M}_0([a, b]; \mathbb{R})$ y sean c_1, c_2 dos números reales entonces $c_1 g_1 + c_2 g_2 \in \mathcal{M}_0([a, b]; \mathbb{R})$ y

$$\int_a^b [c_1 g_1(t) + c_2 g_2(t)] dB_t = c_1 \int_a^b g_1(t)dB_t + c_2 \int_a^b g_2(t)dB_t$$

Lema 2.2.2. Para cualquier $f \in \mathcal{M}^2([a, b]; \mathbb{R})$, existe una sucesión $\{g_n\}$ de procesos simples tales que $\lim_{n \rightarrow \infty} E[\int_a^b |f(t) - g_n(t)|^2 dt] = 0$

La segunda definición de integral de Itô, que se presentará a continuación hace referencia a una extensión de la integral de Itô de procesos simples $g \in \mathcal{M}_0([a, b]; \mathbb{R})$ a procesos más generales $g \in \mathcal{M}^2([a, b]; \mathbb{R})$ y está basada en la siguiente aproximación.

Definición 2.2.5 (Integral de Itô parte II). Sea $f \in \mathcal{M}^2([a, b]; \mathbb{R})$. La integral de Itô de f con respecto a $\{B_t\}$ es definida por

$$\int_a^b f(t)dB_t = \lim_{n \rightarrow \infty} \int_a^b g_n(t)dB_t$$

Donde $\{g_n\}$ es una sucesión de procesos simples tales que

$$\lim_{n \rightarrow \infty} E \left[\int_a^b |f(t) - g_n(t)|^2 dt \right] = 0$$

En la definición 2.2.5 el límite se mantiene independiente de la sucesión particular $\{g_n\}$. Por ejemplo, si $\{h_n\}$ es otra sucesión de procesos simples que converge a f en el sentido que $\lim_{n \rightarrow \infty} E \left[\int_a^b |f(t) - h_n(t)|^2 dt \right] = 0$, entonces la sucesión $\{\varphi_n\}$, donde $\varphi_{2n-1} = g_n$ y $\varphi_{2n} = h_n$ son también convergentes a f en el mismo sentido.

Teorema 2.2.3. Sean $f, g \in \mathcal{M}^2([a, b]; \mathbb{R})$ y sean α, β dos números reales, entonces

$$\text{I) } \int_a^b f(t)dB_t \text{ es } \mathfrak{F}_b\text{-medible}$$

$$\text{II) } E \int_a^b f(t)dB_t = 0$$

$$\text{III) } E \left| \int_a^b f(t)dB_t \right|^2 = E \int_a^b |f(t)|^2 dt$$

$$\text{IV) } \int_a^b [\alpha f(t) + \beta g(t)] dB_t = \alpha \int_a^b f(t)dB_t + \beta \int_a^b g(t)dB_t$$

Teorema 2.2.4. Sea $f \in \mathcal{M}^2([a, b]; \mathbb{R})$. Entonces $E \left[\int_a^b f(t)dB_t / \mathfrak{F}_a \right] = 0$:

$$\begin{aligned} E \left[\left| \int_a^b f(t)dB_t \right|^2 / \mathfrak{F}_a \right] &= E \left[\int_a^b |f(t)|^2 dt / \mathfrak{F}_a \right] \\ &= \int_a^b E [|f(t)|^2 / \mathfrak{F}_a] dt. \end{aligned}$$

Lema 2.2.5. Si $f \in \mathcal{M}^2([a, b]; \mathbb{R})$ y ξ es una variable aleatoria \mathfrak{F}_a -medible de valor real, entonces $\xi f \in \mathcal{M}^2([a, b]; \mathbb{R})$ y $\int_a^b \xi f(t)dB_t = \xi \int_a^b f(t)dB_t$.

Es claro que si $f \in \mathcal{M}^2([a, b]; \mathbb{R})$ entonces $\xi f \in \mathcal{M}^2([a, b]; \mathbb{R})$. Si f es un proceso simple entonces la igualdad anterior es inmediata por la definición de integral estocástica. Para el caso general en el que $f \in \mathcal{M}^2([a, b]; \mathbb{R})$, se toma $\{g_n\}$ como una sucesión de procesos simples que satisfacen $\lim_{n \rightarrow \infty} E \int_a^b |\xi f(t) - \xi g_n(t)|^2 dt = 0$. De igual modo la igualdad se mantiene y la variable aleatoria ξ actúa como una constante.

Definición 2.2.6. Sea $f \in \mathcal{M}([0, T]; \mathbb{R})$. Se define $I(t) = \int_0^t f(s)dB_s$ para $0 \leq t \leq T$, donde $I(0) = \int_0^0 f(s)dB_s = 0$. $I(t)$ se conoce con el nombre de *integral indefinida de Itô sobre f* .

2.2.1. La Fórmula de Itô

Cuando se intenta evaluar una integral estocástica, no es muy apropiado utilizar la definición rigurosa. Esta situación es similar al caso de evaluar una integral ordinaria realizando las sumatorias de Riemann, en lugar de utilizar el teorema fundamental del cálculo y la regla de la cadena.

A diferencia del cálculo ordinario, que comprende teoría de diferenciación e integración, este contexto solo posee teoría de integración; sin embargo resulta posible establecer una versión de la regla de la cadena para integrales de Itô, la cual se conoce como la **Fórmula de Itô**.

Sea $\{B_t\}_{t \geq 0}$ un Movimiento Browniano Unidimensional definido sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, p)$ adaptado a la filtración $\{\mathfrak{F}_t\}_{t \geq 0}$. $\mathcal{L}^1(\mathbb{R}_+, \mathbb{R}^n)$ denota la familia de todos los procesos $f = \{f(t)\}_{t \geq 0}$ que son $\{\mathfrak{F}_t\}$ -adaptados medibles, de valores en \mathbb{R}^n tales que $\int_0^T |f(t)|dt < \infty$ para toda $T > 0$.

Definición 2.2.7. Un proceso de Itô unidimensional es un proceso $\{x(t)\}_{t \geq 0}$ adaptado continuo de la forma $x(t) = x(0) + \int_0^t f(s)ds + \int_0^t g(x)dB_s$, donde $f \in \mathcal{L}^1(\mathbb{R}_+, \mathbb{R})$ y $g \in \mathcal{L}^2(\mathbb{R}_+, \mathbb{R})$. Se dice que $x(t)$ tiene *diferencial estocástica* $dx(t); t \geq 0$ dada por $dx(t) = f(t)dt + g(t)dB_t$.

$C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+; \mathbb{R})$ denota la familia de todas las funciones $V(x, t)$ de valor real definidas sobre $\mathbb{R}^n \times \mathbb{R}$ que son al menos dos veces continuamente diferenciables en x y una vez en t . Si $V \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+; \mathbb{R})$ se tomará

$$V_t = \frac{\partial V}{\partial t} \quad V_x = \left(\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \dots, \frac{\partial V}{\partial x_n} \right);$$

$$V_{xx} = \left(\frac{\partial^2 V}{\partial x_i \partial x_j} \right)_{n \times n} = \begin{pmatrix} \frac{\partial^2 V}{\partial x_1^2} & \cdots & \frac{\partial^2 V}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 V}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 V}{\partial x_n^2} \end{pmatrix}$$

Teorema 2.2.6 (Fórmula de Itô Unidimensional). *Sea $\{x(t)\}_{t \geq 0}$ un proceso de Itô con diferencial estocástica $dx(t) = f(t)dt + g(t)dB_t$, donde $f \in \mathcal{L}^1(\mathbb{R}_+; \mathbb{R})$ y $g \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R})$. Se tomará $V \in C^{2,1}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$. Entonces $V(x(t), t)$ define un nuevo proceso de Itô con diferencial estocástica dada por*

$$dV(x(t), t) = \left[V_t(x(t), t) + V_x(x(t), t)f(t) + \frac{1}{2}V_{xx}(x(t), t)g^2(t) \right] dt + V_x(x(t), t)g(t)dB_t.$$

La fórmula de Itô puede extenderse a una versión Multidimensional, pero antes se debe definir el Movimiento Browniano Multidimensional.

El proceso estocástico $B(t) = (B_1(t), B_2(t), \dots, B_m(t))^T, t \geq 0$ se conoce como *Movimiento Browniano m -dimensional* y es definido sobre un espacio de probabilidad completo $(\Omega, \mathfrak{F}, p)$ adaptado a la filtración $\{\mathfrak{F}_t\}_{t \geq 0}$.

En general $\mathcal{L}^p([a, b]; \mathbb{R}^n)$ denota la familia de todos los procesos $f = \{f(t)\}_{a \leq t \leq b}$ que son $\{\mathfrak{F}_t\}$ -adaptados medibles, de valores en \mathbb{R}^n tales que $\int_a^b |f(t)|^p dt < \infty$.

Definición 2.2.8. Un proceso de Itô n -dimensional es un proceso $x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T; t \geq 0$ adaptado, continuo y con valores en \mathbb{R}^n de la forma

$$x(t) = x(0) + \int_0^t f(s)ds + \int_0^t g(s)dB(s),$$

donde $f = (f_1, f_2, \dots, f_n)^T \in \mathcal{L}^1(\mathbb{R}_+; \mathbb{R}^n)$ y $g = (g_{ij})_{n \times m} \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{n \times m})$. Se dice que $x(t)$ tiene diferencial estocástica $dx(t); t \geq 0$ dada por $dx(t) = f(t)dt + g(t)dB_t$ donde $B(t)$ es un Movimiento Browniano m -dimensional.

Teorema 2.2.7 (Fórmula de Itô Multidimensional). *Sea $x(t)$ en proceso de Itô n -dimensional con $t \geq 0$ con diferencial estocástico dado por $dx(t) = f(t)dt + g(t)dB(t)$, donde $f \in \mathcal{L}^1(\mathbb{R}_+; \mathbb{R}^n)$ y $g \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{n \times m})$. Considere $V \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+; \mathbb{R})$ entonces $V(x(t), t)$ define un nuevo proceso de Itô con diferencial estocástica dada por*

$$dV(x(t), t) = \left[V_t(x(t), t) + V_x(x(t), t)f(t) + \frac{1}{2} \text{traza} (g^T(t)V_{xx}(x(t), t)g(t)) \right] dt + V_x(x(t), t)g(t)dB(t) \quad (2.2)$$

En la ecuación 2.2 se introduce la tabla de multiplicación formal:

$$dt \cdot dt = 0; \quad dB_i \cdot dt = 0; \quad dB_i \cdot dB_i = dt; \quad dB_i \cdot dB_j = 0 \text{ si } i \neq j.$$

Las integrales estocásticas también tienen una versión de integración por partes análoga a la integración por partes del cálculo ordinario. Se verá a continuación esta versión:

Teorema 2.2.8 (Fórmula de Integración por partes). *Sea $\{x(t)\}_{t \geq 0}$ un proceso de Itô Unidimensional con diferencial estocástico*

$$dx(t) = f(t)dt + g(t)dB_t$$

donde $f \in \mathcal{L}^1(\mathbb{R}_+; \mathbb{R})$ y $g \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{1 \times m})$. Sea $\{y(t)\}_{t \geq 0}$ un proceso adaptado continuo de valor real con variación finita. Entonces,

$$d[x(t)y(t)] = y(t)dx(t) + x(t)dy(t),$$

en otras palabras, $x(t)y(t) - x(0)y(0) = \int_0^t y(s)[f(s)ds + g(s)dB_s] + \int_0^t x(s)dy(s)$, donde esta última integral se conoce como la integral de Lebesgue-Stieltjes.

Capítulo 3

Ecuaciones Diferenciales Estocásticas

Las Ecuaciones Diferenciales Estocásticas (E.D.E) tienen múltiples aplicaciones en Matemáticas, Física, Demografía y Finanzas. Estas tienen un tratamiento similar a las Ecuaciones Diferenciales Ordinarias en el momento de hallar soluciones explícitas.

En este capítulo se presentarán algunos conceptos básicos y los teoremas de existencia y unicidad. También se hará la descripción de algunos métodos de solución, se darán algunas soluciones explícitas para Ecuaciones Diferenciales Estocásticas Lineales y se presentarán los métodos de aproximación numérica de Euler y Milstein.

Sea $(\Omega, \mathfrak{F}, p)$ un espacio de probabilidad completo con una filtración $\mathfrak{F}_t \geq 0$.

Sea $B(t) = (B_1(t), B_2(t), \dots, B_m(t))^T, t \geq 0$ un Movimiento Browniano m-dimensional definido sobre este espacio.

Tome $0 \leq t_0 < T < \infty$.

Sea x_0 una variable aleatoria de valores en \mathbb{R}^n \mathfrak{F}_{t_0} -medible tal que $E|x_0|^2 < \infty$.

Sea $f : \mathbb{R}^n \times [t_0, T] \rightarrow \mathbb{R}^n$ y $g : \mathbb{R}^n \times [t_0, T] \rightarrow \mathbb{R}^{n \times m}$ ambas Borel-medibles.

Considere la Ecuación Diferencial Estocástica n-dimensional del tipo

$$dx(t) = f(x(t), t)dt + g(x(t), t)dB(t); t_0 \leq t < T \quad (3.1)$$

Con valor inicial $x(t_0) = x_0$. Por definición de diferencial estocástica, esta ecuación es equivalente a la ecuación integral estocástica dada por:

$$x(t) = x_0 + \int_{t_0}^t f(x(s), s)ds + \int_{t_0}^t g(x(s), s)dB(s); t_0 \leq t \leq T \quad (3.2)$$

Definición 3.0.9. Un proceso estocástico $\{x(t)\}_{t_0 \leq t \leq T}$ con valores en \mathbb{R}^n se conoce como solución de la ecuación (3.1) si tiene las siguientes propiedades:

- 1) $\{x(t)\}$ es continua y \mathfrak{F}_t adaptada

II) $\{f(x(t), t)\} \in \mathcal{L}^1([t_0, T]; \mathbb{R}^n)$ y $\{g(x(t), t)\} \in \mathcal{L}^2([t_0, T]; \mathbb{R}^{n \times m})$

III) La ecuación (3.2) se cumple para todo $t \in [t_0, T]$, con probabilidad 1.

Una solución $\{x(t)\}$ se dice es única si cualquier otra solución $\{\bar{x}(t)\}$ es idéntica a $\{x(t)\}$, es decir:

$$P\{x(t) = \bar{x}(t)\} = 1; \quad t \in [t_0, T]$$

La solución de la ecuación (3.1) puede denotarse por $x(t; t_0, x_0)$. Además de la ecuación (3.2) es claro que para toda $S \in [t_0, T]$,

$$x(t) = x(S) + \int_S^t f(x(r), r)dr + \int_S^t g(x(r), r)dB(r); \quad S \leq t \leq T$$

Pero esta última ecuación es una Ecuación Diferencial Estocástica sobre $[S, T]$ con valor inicial $x(S) = x(S; t_0, x_0)$, cuya solución puede escribirse como $x(t; S, x(S; t_0, x_0))$. De este modo puede observarse que la solución de la ecuación (3.1) satisface la propiedad

$$x(t; t_0, x_0) = x(t; S, x(S; t_0, x_0)); \quad t_0 \leq S \leq t \leq T$$

3.1. Existencia y Unicidad de las Soluciones

Teorema 3.1.1. *Suponga que existen dos constantes positivas K y \bar{K} tales que:*

1. Condición de Lipschitz: *Para toda $x, y \in \mathbb{R}^n$ y $t \in [t_0, T]$*

$$|f(x, t) - f(y, t)|^2 \vee |g(x, t) - g(y, t)|^2 \leq \bar{K} |x - y|^2$$

2. Condición Lineal de Crecimiento: *Para toda $(x, t) \in \mathbb{R}^n \times [t_0, T]$*

$$|f(x, t)|^2 \vee |g(x, t)|^2 \leq K(1 + |x|^2)$$

Entonces existe una única solución $x(t)$ a la ecuación (3.1) y la solución pertenece a $\mathcal{M}^2([t_0, T]; \mathbb{R}^n)$.

Teorema 3.1.2. *Suponga que la condición de crecimiento lineal se cumple, pero la condición de Lipschitz se reemplaza por la siguiente **Condición de Lipschitz Local**:*

Para cada entero $n \geq 1$, existe una constante positiva K_n tal que para toda $t \in [t_0, T]$ y toda $x, y \in \mathbb{R}^n$ con $|x| \vee |y| \leq n$,

$$|f(x, t) - f(y, t)|^2 \vee |g(x, t) - g(y, t)|^2 \leq K_n |x - y|^2$$

Entonces existe una solución única $x(t)$ a la ecuación (3.1) en $\mathcal{M}^2([t_0, T]; \mathbb{R}^n)$.

Teorema 3.1.3. *Suponga que la condición de Lipschitz local se cumple, pero la condición de crecimiento lineal se reemplaza por la siguiente **Condición Monótona**: Existe una constante positiva K tal que para toda $(x, y) \in \mathbb{R}^n \times [t_0, T]$*

$$x^T f(x, t) + \frac{1}{2} |g(x, t)|^2 \leq K(1 + |x|^2)$$

Entonces existe una única solución $x(t)$ a la ecuación (3.1) en $\mathcal{M}^2([t_0, T]; \mathbb{R}^n)$

La condición local de Lipschitz garantiza que la solución existe en $[t_0, \tau_\infty]$, donde $\tau_\infty = \lim_{n \rightarrow \infty} \tau_n$. Pero la condición monótona en lugar de la condición de crecimiento lineal, garantiza que $\tau_\infty = T$, es decir, que la solución existe sobre el intervalo $[t_0, T]$.

Considere la E.D.E dada por

$$dx(t) = f(x(t), t)dt + g(x(t), t)dB(t) \quad (3.3)$$

con valor inicial $x(T_0) = x_0$. Si las suposiciones de existencia y unicidad se cumplen para cada subintervalo finito $[t_0, T]$ de $[t_0, \infty)$, entonces la ecuación (3.3) tiene una solución única $x(t)$ sobre el intervalo completo $[t_0, \infty]$. Esta solución, se llama **Solución Global**.

Teorema 3.1.4. *Suponga que para cada número real $T > t_0$ y para cada entero $n \geq 1$, existe una constante positiva $K_{T,n}$ tal que para toda $t \in [t_0, T]$ y toda $x, y \in \mathbb{R}^n$ con $|x| \vee |y| \leq n$, se tiene $|f(x, t) - f(y, t)|^2 \vee |g(x, t) - g(y, t)|^2 \leq K_{T,n} |x - y|^2$. Suponga además que para toda $T > t_0$, existe una constante K_T positiva tal que para toda $(x, t) \in \mathbb{R}^n \times [t_0, T]$,*

$$x^T f(x, t) + \frac{1}{2} |g(x, t)|^2 \leq K_T(1 + |x|^2)$$

Entonces existe una única solución global $x(t)$ de la ecuación (3.3) y la solución pertenece a $\mathcal{M}^2([t_0, \infty]; \mathbb{R}^n)$.

3.2. Ecuaciones Diferenciales Estocásticas Lineales

En general, las Ecuaciones Diferenciales Estocásticas No Lineales no tienen solución explícita y en la práctica, pueden utilizarse soluciones aproximadas. Sin embargo, es posible encontrar soluciones explícitas para las ecuaciones lineales.

En esta sección, se obtendrá la solución explícita de la Ecuación Diferencial Estocástica General n-dimensional

$$dx(t) = (F(t)x(t))dt + \sum_{k=1}^m (G_k(t)x(t) + g_k(t))dB_k(t) \quad (3.4)$$

sobre $[t_0, T]$, donde $F(\cdot)$, $G_k(\cdot)$ son funciones de valor matricial de orden $n \times n$, $f(\cdot)$, $g_k(\cdot)$ son funciones de valores en \mathbb{R}^n y $B(t) = (B_1(t), B_2(t), \dots, B_m(t))^T$ es un Movimiento

Browniano m -dimensional.

La ecuación lineal se dice es homogénea si $f(t) = g(t) = \dots \equiv 0$; se dice que es lineal en el sentido estrecho si $G_1(t) = G_2(t) = \dots \equiv 0$ y se dice que es autónoma si los coeficientes F, f, G_k, g_k son todos independientes de t .

También se asumirá que F, f, G_k, g_k son todos Borel-medibles y acotados sobre $[t_0, T]$. De este modo, por el Teorema de existencia y unicidad, la ecuación (3.4) tiene una única solución continua en $\mathcal{M}^2([t_0, T]; \mathbb{R}^n)$ para cada valor inicial $x(t_0) = x_0$, el cual es \mathfrak{F}_{t_0} -medible y pertenece a $\mathcal{L}^2(\Omega; \mathbb{R}^n)$.

3.2.1. Fórmula Estocástica de Liouville

Considere la Ecuación Diferencial Estocástica Lineal Homogénea

$$dx(t) = (F(t)x(t))dt + \sum_{k=1}^m G_k(t)x(t)dB_k(t); t \in [t_0, T] \quad (3.5)$$

Según lo asumido $F(t) = (F_{ij}(t))_{n \times n}$, $G_k = (G_{ij}^k(t))_{n \times n}$ son todas Borel-medibles y acotadas. Para cada $j = 1, 2, \dots, n$, tiene e_j como un vector columna unitario en la dirección de x_j , es decir, $e_j = \underbrace{(0, 0, \dots, 1, 0, \dots, 0)}_j^T$.

Sea $\Phi_j(t) = (\Phi_{1j}(t), \Phi_{2j}(t), \dots, \Phi_{nj}(t))^T$ la solución de la ecuación (3.5) con valor inicial $x(t_0) = e_j$. Defina la matriz de dimensión $n \times n$

$$\Phi(t) = (\Phi_1(t), \Phi_2(t), \dots, \Phi_n(t)) = (\Phi_{ij}(t))_{n \times n}$$

$\Phi(t)$ se conoce como matriz fundamental de la ecuación (3.5). Observe que $\Phi(t_0)$ es la matriz identidad de orden $n \times n$ y

$$d\Phi(t) = F(t)\Phi(t)dt + \sum_{k=1}^m G_k(t)\Phi(t)dB_k(t) \quad (3.6)$$

La ecuación (3.6) también puede expresarse de la siguiente forma:

Para $i \leq 1, j \leq n$,

$$d\Phi_{ij}(t) = \sum_{l=1}^n F_{il}(t)\Phi_{lj}(t)dt + \sum_{k=1}^m \sum_{l=1}^n G_{il}^k(t)\Phi_{lj}(t)dB_k(t) \quad (3.7)$$

El siguiente teorema muestra que cualquier solución de la ecuación (3.1) puede ser expresada en términos de $\Phi(t)$ y esta es la razón por la cual $\Phi(t)$ se conoce como matriz fundamental.

Teorema 3.2.1. *Dada la condición inicial $x(t_0) = x_0$, la única solución para la ecuación (3.1) está dada por $x(t) = \Phi(t)x_0$*

Definición 3.2.1. Se denota por $W(t)$ al determinante de la matriz fundamental, $\Phi(t)$, es decir, $W(t) = \det(\Phi(t))$

$W(t)$ se conoce como el determinante Wroskiano Estocástico, donde $W(t_0) = 1$

Teorema 3.2.2 (Fórmula de Liouville Estocástica). *El determinante Wroskiano Estocástico $W(t)$ está dado por la expresión*

$$W(t) = \exp \left[\left(\int_{t_0}^t \text{tr}(F(s)) - \frac{1}{2} \sum_{k=1}^m \text{tr}(G_k(s))G_k^T(s) \right) ds + \sum_{k=1}^m \int_{t_0}^t \text{tr}(G_k(s))dB_k(s) \right] \quad (3.8)$$

Lema 3.2.3. *Sean $a(\cdot)$ y $b_k(\cdot)$ funciones acotadas Borel-medibles de valor real sobre $[t_0, T]$, luego*

$$y(t) = y_0 \exp \left[\int_{t_0}^t \left(a(s) - \frac{1}{2} \sum_{k=1}^m b_k^2(s) \right) ds + \sum_{k=1}^m \int_{t_0}^t b_k(s)dB_k(s) \right] \quad (3.9)$$

es la única solución de la Ecuación Diferencial Estocástica Lineal Escalar

$$dy(t) = a(t)y(t)dt + \sum_{k=1}^m b_k(t)y(t)dB_k(t); \quad t \in [t_0, T] \quad (3.10)$$

con valor inicial $y(t_0) = y_0$.

La Fórmula de Liouville Estocástica (3.8), implica directamente que $W(t) > 0$, es decir, para toda $t \in [t_0, T]$, la cual a su vez implica que $\Phi(t)$ es invertible como se establece en el siguiente teorema.

Teorema 3.2.4. *Para toda $t \in [t_0, T]$, la matriz fundamental $\Phi(t)$ es invertible con probabilidad 1. La matriz $\Phi^{-1}(t)$ se llama la inversa de la matriz $\Phi(t)$.*

3.2.2. Fórmula de Variación de las Constantes

Considere de nuevo la Ecuación Diferencial Estocástica Lineal General n-dimensional

$$dx(t) = (F(t)x(t) + f(t))dt + \sum_{k=1}^m (G_k(t)x(t) + g_k(t))dB_k(t) \quad (3.11)$$

donde $t \in [t_0, T]$ con valor inicial $x(t_0) = x_0$. La ecuación

$$dx(t) = (F(t)x(t))dt + \sum_{k=1}^m (G_k(t)x(t) + g_k(t))dB_k(t) \quad (3.12)$$

se llama Ecuación Homogénea correspondiente del primer sistema (3.11). En esta sección se establecerá una fórmula muy útil, que se conoce como **Fórmula de Variación de las Constantes**, la cual representa la única solución para la ecuación (3.11) en términos de la matriz fundamental para la correspondiente ecuación homogénea (3.12).

Teorema 3.2.5. *La única solución para la ecuación (3.11) puede ser expresada como*

$$x(t) = \Phi(t) \left(x_0 + \int_{t_0}^t \Phi^{-1}(t) \left[f(s) - \sum_{k=1}^m G_k(s)g_k(s) \right] ds + \sum_{k=1}^m \int_{t_0}^t \Phi^{-1}(t)g_k(s)dB_k(s) \right)$$

donde $\Phi(t)$ es la matriz fundamental de la ecuación homogénea correspondiente.

A continuación se presentarán 3 tipos diferentes de Ecuaciones Diferenciales Estocásticas en las que la matriz fundamental $\Phi(t)$ es de gran importancia para hallar la solución explícita de cada ecuación bajo la condición inicial $x(t_0) = x_0$ sobre el intervalo $[t_0, T]$.

3.2.3. Ecuaciones Lineales Escalares

Considere la Ecuación Diferencial Estocástica Lineal Escalar General dada por

$$dx(t) = (a(t)x(t) + \bar{a}(t))dt + \sum_{k=1}^m (b_k(t)x(t) + \bar{b}_k(t))dB_k(t) \quad (3.13)$$

sobre $[t_0, T]$ con valor inicial $x(t_0) = x_0$. En este caso $x_0 \in \mathcal{L}^2(\Omega; \mathbb{R})$ es \mathfrak{F}_{t_0} -medible, y $a(t), \bar{a}(t), b_k(t), \bar{b}_k(t)$ son funciones escalares acotadas Borel-medibles sobre $[t_0, T]$. La ecuación lineal homogénea correspondiente está dada por

$$dx(t) = a(t)x(t)dt + \sum_{k=1}^m b_k(t)x(t)dB_k(t) \quad (3.14)$$

Por el lema 3.2.3, la solución fundamental para la ecuación (3.14) está dada por

$$\Phi(t) = \exp \left[\int_{t_0}^t \left(a(s) - \frac{1}{2} \sum_{k=1}^m b_k^2(s) \right) ds + \sum_{k=1}^m \int_{t_0}^t b_k(s)dB_k(s) \right].$$

Ahora, aplicando el teorema 3.2.5, se obtiene la solución explícita para la ecuación (3.13) dada por

$$x(t) = \Phi(t) \left(x_0 + \int_{t_0}^t \Phi^{-1}(t) \left[\bar{a}(s) - \sum_{k=1}^m b_k(s)\bar{b}_k(s) \right] ds + \sum_{k=1}^m \int_{t_0}^t \Phi^{-1}(t)\bar{b}_k(s)dB_k(s) \right).$$

3.2.4. Ecuaciones Lineales en el Sentido Estricto

Considere la Ecuación Diferencial Estocástica Lineal n -dimensional dada por

$$dx(t) = (F(t)x(t) + f(t))dt + \sum_{k=1}^m g_k(t)dB_k(t) \quad (3.15)$$

sobre $[t_0, T]$ con valor inicial $x(t_0) = x_0$, donde F, f, g_k son Borel-medibles y acotadas sobre $[t_0, T]$ y $x(t_0) = x_0$ es \mathfrak{F}_{t_0} -medible y pertenece a $\mathcal{L}^2(\Omega; \mathbb{R}^n)$.

La ecuación lineal homogénea correspondiente es la ecuación diferencial ordinaria

$$dx(t) = F(t)x(t)dt \quad (3.16)$$

La solución para la ecuación (3.15) tiene la forma

$$x(t) = \Phi(t) \left(x_0 + \int_{t_0}^t \Phi^{-1}(s)f(s)ds + \sum_{k=1}^m \int_{t_0}^t \Phi^{-1}(s)g_k(s)dB_k(s) \right) \quad (3.17)$$

donde $\Phi(t)$ es la matriz fundamental para la ecuación (3.16).

En particular, cuando $F(t)$ es independiente de t , es decir, $F(t) = F$ es una matriz constante de dimensión $n \times n$, la matriz fundamental $\Phi(t)$ tiene la forma simple $\Phi(t) = e^{F(t-t_0)}$ y su matriz inversa $\Phi^{-1}(t) = e^{-F(t-t_0)}$, la ecuación (3.15) tiene la solución explícita dada por

$$x(t) = e^{F(t-t_0)} \left(x_0 + \int_{t_0}^t e^{-F(t-t_0)} f(s)ds + \sum_{k=1}^m \int_{t_0}^t e^{-F(t-t_0)} g_k(s)dB_k(s) \right)$$

3.2.5. Ecuaciones Lineales Autónomas

Considere la Ecuación Diferencial Estocástica Lineal n -dimensional dada por

$$dx(t) = (F \cdot x(t) + f)dt + \sum_{k=1}^m (G_k \cdot x(t) + g_k)dB_k(s) \quad (3.18)$$

sobre $[t_0, T]$ con la condición inicial $x(t_0) = x_0$, donde $F, G_k \in \mathbb{R}^{n \times n}$ y $f, g_k \in \mathbb{R}^n$.

La ecuación homogénea correspondiente está dada por

$$dx(t) = Fx(t)dt + \sum_{k=1}^m G_k x(t)dB_k(t) \quad (3.19)$$

En general, la matriz fundamental $\Phi(t)$ no es dada en forma explícita. Sin embargo, si las matrices F, G_1, G_2, \dots, G_m conmutan, es decir, si $FG_k = G_kF$, $G_kG_j = G_jG_k$ para toda

$k \geq 1, j \leq m$, entonces la matriz fundamental para la ecuación (3.19) tiene la forma explícita

$$\Phi(t) = \exp \left[\left(F - \frac{1}{2} \sum_{k=1}^m G_k^2 \right) (t - t_0) + \sum_{k=1}^m G_k (B_k(t) - B_k(t_0)) \right]$$

de este modo la **Ecuación Lineal Autónoma** (3.18) tiene solución explícita dada por

$$x(t) = \Phi(t) \left[x_0 + \left(\int_{t_0}^t \Phi^{-1}(s) ds \right) \left(f - \sum_{k=1}^m G_k g_k \right) + \sum_{k=1}^m \left(\int_{t_0}^t \Phi^{-1}(s) dB_k(s) \right) g_k \right]$$

3.3. Aproximaciones Numéricas para EDE

Muchas Ecuaciones Diferenciales Estocásticas no pueden resolverse explícitamente, por esto es conveniente disponer de métodos numéricos que permitan la simulación de soluciones. En esta sección se presentarán los métodos de aproximación de Euler y Milstein con los coeficientes de error que determinan la medida de la calidad en la aproximación, basados en la diferencia entre los momentos de primer orden y en la diferencia entre la solución y su aproximación, y además se dará un ejemplo de cada método de aproximación para el caso unidimensional.

3.3.1. Método de Aproximación de Euler

El método de aproximación de Euler⁴ para Ecuaciones Diferenciales Estocásticas es similar al método de Euler para resolver Ecuaciones Diferenciales ordinarias, y está basado en la discretización de dos integrales y su respectiva aproximación dada por la evaluación del integrando en el punto inferior de cada subintervalo $[t_i, t_{i+1}]$ sobre $[0, T]$.

Considere la Ecuación Diferencial Estocástica dada por

$$dX_t = f(X_t)dt + g(X_t)dB_t \quad (3.20)$$

sobre $[0, T]$ con la condición inicial $X(0) = X_0$, donde f y g son dos funciones escalares.

La ecuación (3.20) puede expresarse en forma integral como

$$X_t = X_0 + \int_0^t f(X_s)ds + \int_0^t g(X_s)dB_s; \quad 0 \leq t \leq T.$$

Tome la subdivisión τ_n del intervalo $[0, T]$ dada por $\tau_n : 0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T$.

Luego, la solución X_t de la ecuación diferencial será aproximada en cada punto t_i de la partición por

$$X_{t_i} = X_{t_{i-1}} + \int_{t_{i-1}}^{t_i} f(X_s)ds + \int_{t_{i-1}}^{t_i} g(X_s)dB_s; \quad i = 1, 2, \dots, n. \quad (3.21)$$

La aproximación de Euler se basa en la discretización de esta última ecuación, por medio de las siguientes aproximaciones:

$$\text{I) } \int_{t_{i-1}}^{t_i} f(X_s) ds \approx f(X_{t_{i-1}}) \Delta t_i$$

$$\text{II) } \int_{t_{i-1}}^{t_i} g(X_s) dB_s \approx g(X_{t_{i-1}}) \Delta B_i,$$

donde $\Delta t_i = t_i - t_{i-1}$ y $\Delta B_i = B_{t_i} - B_{t_{i-1}}$.

Si se denota por $X_{t_i}^{(n)}$ a la solución aproximada de la ecuación diferencial, la aproximación de Euler para X_{t_i} está dada por:

$$X_{t_i}^{(n)} = X_{t_{i-1}}^{(n)} + f(X_{t_{i-1}}^{(n)}) \Delta t_i + g(X_{t_{i-1}}^{(n)}) \Delta B_i \quad \text{con } i = 1, 2, \dots, n.$$

El esquema general para el método de Euler está dado por

$$\begin{aligned} X_0^{(n)} &= X_0 \\ X_{t_1}^{(n)} &= X_0^{(n)} + f(X_0^{(n)}) \Delta t + g(X_0^{(n)}) \Delta B_{t_1} \\ X_{t_2}^{(n)} &= X_{t_1}^{(n)} + f(X_{t_1}^{(n)}) \Delta t + g(X_{t_1}^{(n)}) \Delta B_{t_2} \\ &\vdots \\ X_T^{(n)} &= X_{t_{n-1}}^{(n)} + f(X_{t_{n-1}}^{(n)}) \Delta t + g(X_{t_{n-1}}^{(n)}) \Delta B_{t_n}. \end{aligned}$$

Para analizar la calidad de las soluciones aproximadas por el método de Euler y la aproximación de Milstein, la cual será explicada posteriormente, y teniendo en cuenta que X_t y $X_t^{(n)}$ deberán estar “acopladas” hasta el final del intervalo $[0, T]$, se introduce el coeficiente de error e_s , que mide la diferencia entre la solución de la ecuación diferencial estocástica y la solución aproximada dado por $e_s = E \left| X_T - X_T^{(n)} \right|$.

Una buena aproximación de X_t se llama solución numérica fuerte. Para establecer la medida de la calidad de las aproximaciones se introducirán los siguientes conceptos importantes:

Definición 3.3.1. Se dice que $X_t^{(n)}$ es una solución numérica fuerte de la Ecuación Diferencial Estocástica (3.20) si $e_s \rightarrow 0$ cuando $\Delta t \rightarrow 0$.

Definición 3.3.2. Se dice que la solución numérica $X_t^{(n)}$ converge fuertemente a X_t con orden $r > 0$ si existe una constante $c > 0$ tal que $e_s < c \Delta t^r$ para Δt lo suficientemente pequeño.

El orden de convergencia fuerte mide la tasa a la cual la media del error se acerca a cero cuando $\Delta t \rightarrow 0$.

Otra medida de error no está basada en la cercanía de la solución explícita y la solución aproximada, sino en la diferencia de los momentos de primer orden;

$e_w = \left| Ef(X_T) - Ef(X_T^{(n)}) \right|$, donde f es una función suave, generalmente polinómica.

Definición 3.3.3. Se dice que $X_t^{(n)}$ es una solución numérica débil de la Ecuación Diferencial Estocástica (3.20) si $e_w \rightarrow 0$ cuando $\Delta t \rightarrow 0$.

Definición 3.3.4. Se dice que la solución numérica $X_t^{(n)}$ converge débilmente a X_t con orden $r > 0$ si existe una constante $c > 0$ tal que $e_w < c\Delta t^r$ para Δt lo suficientemente pequeño.

El orden de convergencia débil mide la tasa a la cual el error de las medias se acerca a cero cuando $\Delta t \rightarrow 0$.

El método de Euler converge fuertemente con orden $r = 0,5$ y débilmente con orden $r = 1$.

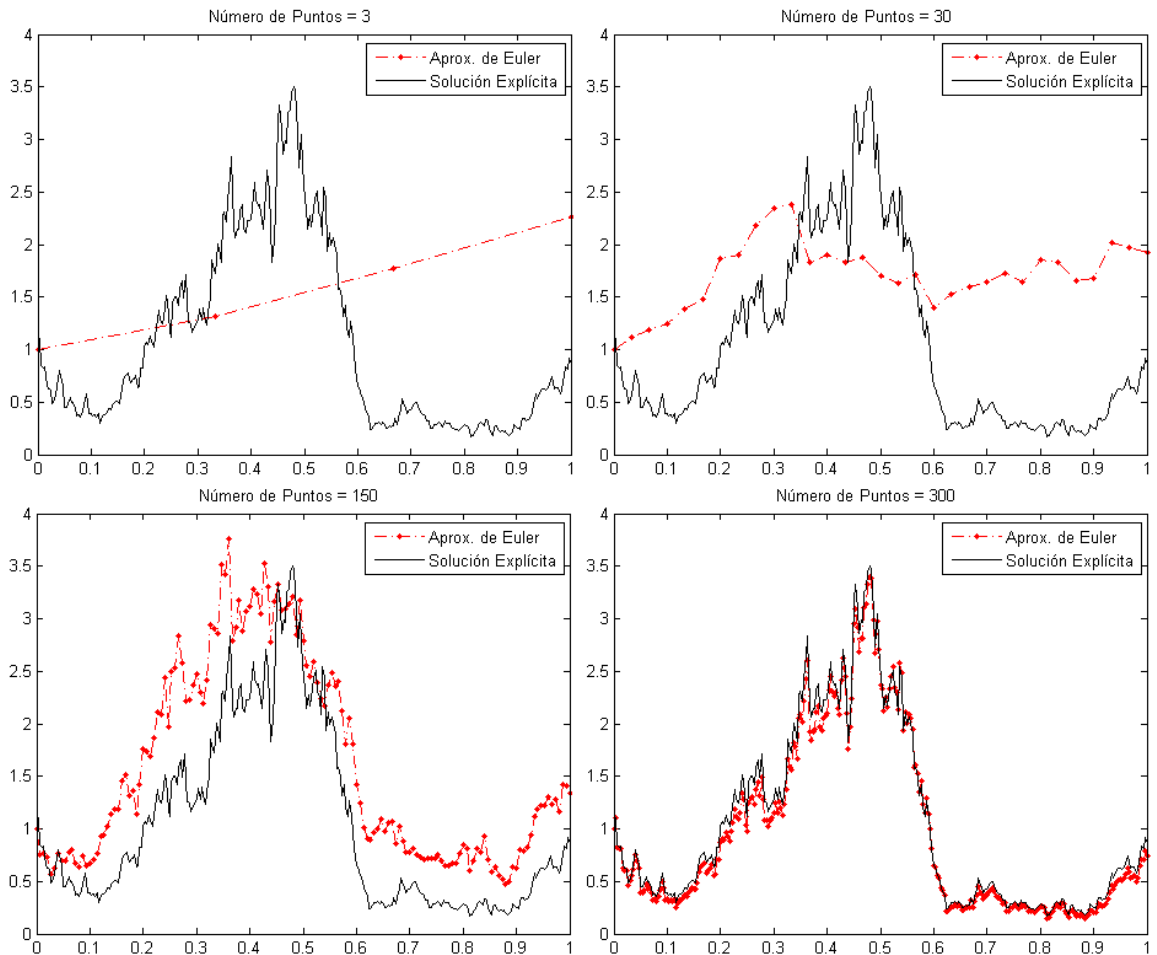


Figura 3.1: Aproximaciones de Euler para $dX_t = \lambda X_t dt + \mu X_t dB_t$ con $\lambda = 1, \mu = 2$

3.3.2. Método de Aproximación de Milstein

La aproximación de Milstein introduce una corrección al método de Euler, que aumenta el orden de convergencia fuerte a 1. La corrección surge porque la expansión tradicional de Taylor debe ser modificada en el sentido de Itô. La idea consiste en aplicar la fórmula de Itô a los integrandos $f(X_s)$ y $g(X_s)$ en la ecuación (3.21) para obtener

$$X_{t_i} - X_{t_{i-1}} = \int_{t_{i-1}}^{t_i} \left[f(X_{t_{i-1}}) + \int_{t_{i-1}}^s (ff' + \frac{1}{2}g^2f'')dy + \int_{t_{i-1}}^{t_i} gf'dB_y \right] ds + \int_{t_{i-1}}^{t_i} \left[g(X_{t_{i-1}}) + \int_{t_{i-1}}^s (ff' + \frac{1}{2}g^2f'')dy + \int_{t_{i-1}}^{t_i} gf'dB_y \right] ds$$

Al considerar la igualdad $\int_{t_{i,1}}^{t_i} \int_{t_{i-1}}^s dB_y dB_s = \frac{1}{2}[(\Delta B_{t_i})^2 - \Delta t_i]$ y la aproximación $\int_{t_{i,1}}^{t_i} \int_{t_{i-1}}^s gg'dB_y dB_s \approx g(X_{t_{i-1}})g'(X_{t_{i-1}}) \int_{t_{i,1}}^{t_i} \int_{t_{i-1}}^s dB_y dB_s$ se obtiene el método de aproximación de Milstein dado por

$$X_{t_i}^{(n)} = X_{t_{i-1}}^{(n)} + f(X_{t_{i-1}}^{(n)})\Delta t + g(X_{t_{i-1}}^{(n)})\Delta B_{t_i} + \frac{1}{2} g(X_{t_{i-1}}^{(n)})g'(X_{t_{i-1}}^{(n)})v [(\Delta B_{t_i})^2 - \Delta t]$$

para $i = 1, 2, \dots, n$.

El esquema general para el método de Milstein está dado por

$$\begin{aligned} X_0^{(n)} &= X_0 \\ X_{t_1}^{(n)} &= X_0^{(n)} + f(X_0^{(n)})\Delta t + g(X_0^{(n)})\Delta B_{t_1} + \frac{1}{2} g(X_0^{(n)})g'(X_0^{(n)})v [(\Delta B_{t_1})^2 - \Delta t] \\ X_{t_2}^{(n)} &= X_{t_1}^{(n)} + f(X_{t_1}^{(n)})\Delta t + g(X_{t_1}^{(n)})\Delta B_{t_2} + \frac{1}{2} g(X_{t_1}^{(n)})g'(X_{t_1}^{(n)})v [(\Delta B_{t_2})^2 - \Delta t] \\ &\vdots \\ X_T^{(n)} &= X_{t_{n-1}}^{(n)} + f(X_{t_{n-1}}^{(n)})\Delta t + g(X_{t_{n-1}}^{(n)})\Delta B_{t_n} + \frac{1}{2} g(X_{t_{n-1}}^{(n)})g'(X_{t_{n-1}}^{(n)})v [(\Delta B_{t_n})^2 - \Delta t]. \end{aligned}$$

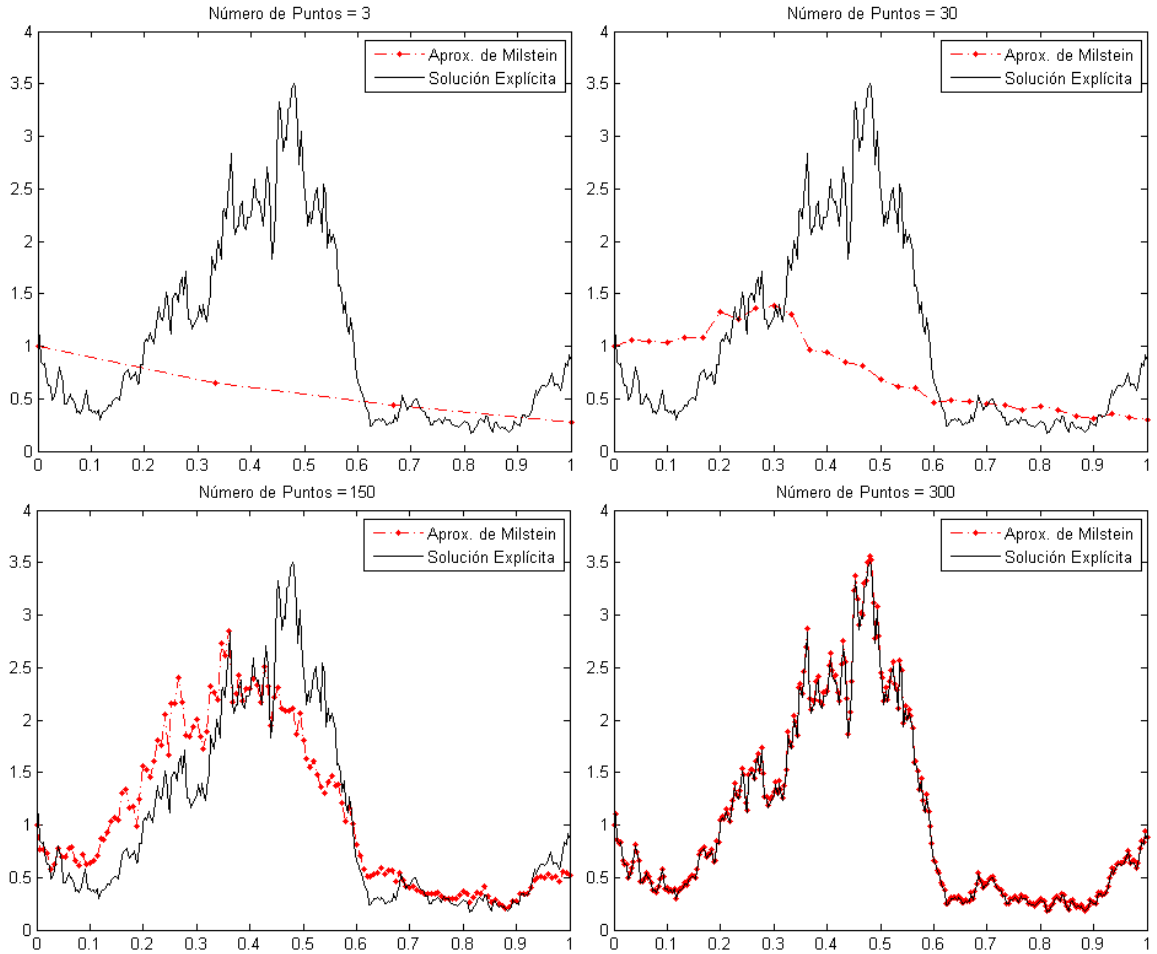


Figura 3.2: Aproximaciones de Milstein para $dX_t = \lambda X_t dt + \mu X_t dB_t$ con $\lambda = 1, \mu = 2$

3.4. Ecuaciones Diferenciales Parciales

En esta sección se explorará la conexión íntima que existe entre las ecuaciones diferenciales estocásticas y ciertas ecuaciones diferenciales parciales parabólicas. También se presentará la fórmula de representación estocástica de Feynman-Kac, que será de gran utilidad más adelante cuando se estudie el modelo de Black-Scholes. Por último se hará una breve descripción de las ecuaciones Backward y Forward de Kolmogorov, y de la ecuación de Fokker-Planck.

Definición 3.4.1. Considere la Ecuación Diferencial Estocástica n -dimensional

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$

El operador diferencial parcial A , conocido como *operador infinitesimal* para X , es definido

para cualquier función $h(x); h \in C^2$ por

$$A \cdot h(t, x) = \sum_{i=1}^n \mu_i(t, x) \frac{\delta h(x)}{\delta x_i} + \frac{1}{2} \sum_{i,j=1}^n C_{ij}(t, x) \frac{\delta^2 h(x)}{\delta x_i \delta x_j},$$

donde $C(t, x) = \sigma(t, x)\sigma'(t, x)$

Este operador también se conoce como operador de Dynkin, operador de Itô, o el operador backward de Kolmogorov. En términos del generador infinitesimal, la fórmula De Itô adquiere la forma

$$df(t, X_t) = \left\{ \frac{\delta f}{\delta t} + \mathcal{A}f \right\} dt + [\nabla_x f] \sigma dW_t$$

donde el gradiente ∇_x está definido para $h \in C^1(\mathbb{R}^n)$ como

$$\nabla_x h = \left[\frac{\delta h}{\delta x_1}, \frac{\delta h}{\delta x_2}, \dots, \frac{\delta h}{\delta x_n} \right].$$

3.4.1. El Problema de Cauchy

Dadas tres funciones escalares $\mu(t, x)$, $\sigma(t, x)$ y $\Phi(x)$, la tarea es encontrar una función F que satisfaga el siguiente problema de valor frontera sobre $[0, T] \times \mathbb{R}$

$$\frac{\partial F}{\partial t}(t, x) + \mu(t, x) \frac{\partial F}{\partial x} + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 F}{\partial x^2} = 0 \quad (3.22)$$

$$F(T, x) = \Phi(x) \quad (3.23)$$

En lugar de atacar este problema usando solamente herramientas analíticas, se construirá una fórmula de representación estocástica, la cual dará la solución a esta ecuación en términos de una EDE que está asociada en forma natural a (3.22) y (3.23). De este modo se hará la suposición que existe una solución F para estas dos ecuaciones.

Tómese un punto fijo t en el tiempo y un punto fijo x en el espacio. Defina el proceso estocástico X_t sobre el intervalo de tiempo $[t, T]$ como la solución a la EDE

$$\begin{aligned} dX_t &= \mu(s, X_s)dt + \sigma(s, X_s)dW_s, \\ X_t &= x, \end{aligned}$$

el generador infinitesimal \mathcal{A} para este operador está dado por

$$\mathcal{A} = \mu(t, x) \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2}{\partial x^2}$$

el cual es exactamente el operador que aparece en la ecuación diferencial parcial (3.22). Así se puede escribir el problema del valor de frontera

$$\begin{aligned}\frac{\partial F}{\partial t}(t, x) + \mathcal{A}F(t, x) &= 0 \\ F(T, x) &= \Phi(x).\end{aligned}\tag{3.24}$$

Aplicando la fórmula de Itô al proceso $F(s, X(s))$ resulta

$$F(T, X_T) = F(t, X_t) + \int_t^T \left[\frac{\partial F}{\partial t}(s, X_s) + \mathcal{A}F(s, X_s) \right] ds + \int_t^T \sigma(s, X_s) \frac{\partial F}{\partial x}(s, X_s) dW_s.$$

Al suponer que F realmente satisface la ecuación (3.24), la primera integral se desvanecerá. Si además el proceso $\mu(s, X_s) \frac{\partial F}{\partial x}(s, X_s)$ es suficientemente integrable y se toman valores esperados, la integral estocástica también se desvanecerá.

El valor inicial $X_t = x$ y la condición de frontera $F(T, x) = \Phi(x)$ finalmente permiten escribir la fórmula

$$F(t, x) = E_{t,x} [\Phi(X_T)],$$

donde se tiene indexado el operador esperanza, para enfatizar el valor esperado es tomado del valor inicial dado $X_t = x$. Así, se ha probado el siguiente resultado, el cual se conoce como la *fórmula de representación estocástica de Feynman-Kac*.

Proposición 3.4.1 (Feynman-Kac I). *Suponga que F es una solución del problema del valor de frontera*

$$\begin{aligned}\frac{\partial F}{\partial t}(t, x) + \mu(t, x) \frac{\partial F}{\partial x} + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 F}{\partial x^2}(t, x) &= 0 \\ F(T, x) &= \Phi(x).\end{aligned}$$

Además, suponga que el proceso

$$\sigma(s, X_s) \frac{\partial F}{\partial x}(s, X_s)$$

está en \mathcal{L}^2 , donde X satisface la EDE

$$\begin{aligned}dX_s &= \mu(s, X_s) dt + \sigma(s, X_s) dW_s \\ X_t &= x.\end{aligned}\tag{3.25}$$

Entonces F tiene una representación

$$F(t, x) = E_{t,x} [\Phi(X_t)].$$

Es claro que se necesita la suposición de integrabilidad $\sigma(s, X_s) \frac{\partial F}{\partial x}(s, X_s) \in \mathcal{L}^2$ para garantizar que el valor esperado de la integral estocástica sea igual a cero. Un problema de valor frontera del tipo anterior, también llamado problema parabólico, tendrá infinitas soluciones (ver John 1982). Sin embargo, la proposición anterior sólo proporcionará la solución “exacta”.

Se puede considerar también el problema del valor de frontera

$$\begin{aligned} \frac{\partial F}{\partial t}(t, x) + \mu(t, x) \frac{\partial F}{\partial x}(t, x) + \sigma^2(t, x) \frac{\partial^2 F}{\partial x^2}(t, x) + rF(t, x) &= 0 \\ F(T, x) &= \Phi(x), \end{aligned}$$

donde r es un número real dado. Las ecuaciones de este tipo aparecen de nuevo en el estudio de los problemas de precios para derivados financieros. Inspirados por la técnica de EDO de factores integrantes, se multiplicará la ecuación completa por el factor e^{rs} , y si se considera el proceso $Z(s) = e^{rs}F(s, X(s))$, donde X es definida como en (3.25), se obtiene el siguiente resultado

Proposición 3.4.2 (Feynman-Kac II). *Suponga que F es una solución del problema del valor de frontera*

$$\begin{aligned} \frac{\partial F}{\partial t}(t, x) + \mu(t, x) \frac{\partial F}{\partial x}(t, x) + \frac{1}{2}\sigma^2(t, x) \frac{\partial^2 F}{\partial x^2}(t, x) + rF(t, x) &= 0 \\ F(T, x) &= \Phi(x). \end{aligned}$$

Suponga además que el proceso $\sigma(s, X_s) \frac{\partial F}{\partial x}(s, X_s) \in \mathcal{L}^2$ donde X es definida más adelante. Entonces F tiene la representación

$$F(t, x) = e^{r(T-t)} E_{t,x} [\Phi(X_t)],$$

donde X satisface la EDE

$$\begin{aligned} dX_s &= \mu(s, X_s)dt + \sigma(s, X_s)dW_s, \\ X_t &= x. \end{aligned}$$

Ejemplo. Resuelva la EDP (Ecuación Diferencial Parcial)

$$\begin{aligned} \frac{\partial F}{\partial t}(t, x) + \frac{1}{2}\sigma^2(t, x) \frac{\partial^2 F}{\partial x^2}(t, x) &= 0, \\ F(T, x) &= x^2, \end{aligned}$$

donde σ es una constante.

Solución:

Por proposición de Feynman-Kac I se tiene que

$$F(t, x) = E_{t,x} [X_T^2], \text{ donde } dX_s = 0ds + \sigma dW_s; X_t = x.$$

Esta ecuación puede ser fácilmente resuelta, por lo tanto $X_t = x + \sigma[W_T - W_t]$; así X_T tiene distribución $N[x, \sigma\sqrt{T-t}]$. En consecuencia se tiene la solución

$$F(t, x) = E[X_T^2] = \text{var}[X_T] + (E[X_T])^2 = \sigma^2(T-t) + x^2.$$

■

Hasta el momento solo se ha tratado el caso escalar, pero con los mismos argumentos anteriores se dará el siguiente resultado.

Proposición 3.4.3. *Suponga que se tienen las siguientes condiciones:*

- Una función (vector columna evaluado) $\mu : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$.
- Una función $C : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow M(n, d)$, la cual puede ser escrita en la forma $C(t, x) = \sigma(t, x)\sigma^\bullet(t, x)$, para alguna función $\sigma : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow M(n, d)$.
- Una función de valores reales $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$.
- Un número real r .

Suponga que $F : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una solución al problema del valor de frontera

$$\frac{\partial F}{\partial t}(t, x) + \sum_{i=1}^n \mu_i(t, x) \frac{\partial F}{\partial x_i}(t, x) + \frac{1}{2} \sum_{i,j=1}^n C_{ij}(t, x) \frac{\partial^2 F}{\partial x_i \partial x_j}(t, x) - rF(t, x) = 0, F(T, x) = \Phi(x).$$

Suponga además que el proceso

$$\sum_{i=1}^n \sigma_i(s, X_s) \frac{\partial F}{\partial x_i}(s, X_s)$$

está en \mathcal{L}^2 , donde X satisface la EDE

$$\begin{aligned} dX_s &= \mu(s, X_s)dt + \sigma(s, X_s)dW_s, \\ X_t &= x. \end{aligned} \tag{3.26}$$

Entonces F tiene la representación

$$F(t, x) = e^{-r(T-t)} E_{t,x} [\Phi(X_T)]. \tag{3.27}$$

Proposición 3.4.4. *Considere un proceso vector X dado con generador \mathcal{A} , y una función $F(t, x)$. Entonces, excepto a de integrabilidad, se cumple lo siguiente:*

- *El proceso $F(t, X_t)$ es una martingala relativa a la filtración \mathfrak{F}^X si y sólo si F satisface la EDP*

$$\frac{\partial F}{\partial t} + \mathcal{A}F = 0.$$

- *Para cada (t, x) y $T \geq t$, se tiene que $F(t, x) = E_{t,x}[F(T, X_T)]$.*

3.4.2. Las Ecuaciones de Kolmogorov

Se usarán algunos de los resultados anteriores para obtener resultados clásicos concernientes a las probabilidades de transición para la solución de algunas Ecuaciones Diferenciales Estocásticas.

Suponga que X es una solución de la ecuación

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \quad (3.28)$$

con generador infinitesimal \mathcal{A} dado por

$$(\mathcal{A}f)(s, y) = \sum_{i=1}^n \mu_i(s, y) \frac{\partial f}{\partial y_i}(s, y) + \frac{1}{2} \sum_{i,j=1}^n C_{ij}(s, y) \frac{\partial^2 f}{\partial y_i \partial y_j}(s, y),$$

donde usualmente $C(t, x) = \sigma(t, x)\sigma^\bullet(t, x)$.

Considere ahora el problema del valor de frontera

$$\begin{aligned} \left(\frac{\partial u}{\partial s} + \mathcal{A}u \right) (s, y) &= 0, \quad (s, y) \in (0, T) \times \mathbb{R}^n, \\ u(T, y) &= I_B(y), \quad y \in \mathbb{R}^n \end{aligned}$$

donde I_B es la función indicadora para el conjunto B . Por la proposición 3.4.3, inmediatamente se tiene

$$u(s, y) = E_{s,y}[I_B(X_T)] = P(X_T \in B | X_s = y),$$

donde X es una solución de la EDE (3.28). Este argumento también puede ser utilizado para probar el siguiente resultado.

Proposición 3.4.5 (Ecuación Backward de Kolmogorov I). *Sea X una solución de la EDE (3.28). Entonces las probabilidades de transición $P(s, y, t, B) = P(X_t \in B | X(s) = y)$ son dadas como la solución a la ecuación*

$$\begin{aligned} \left(\frac{\partial P}{\partial s} + \mathcal{A}P \right) (s, y, t, B) &= 0, \quad (s, y) \in (0, t) \times \mathbb{R}^n, \\ P(s, y, t, B) &= I_B(y). \end{aligned}$$

Usando básicamente el mismo razonamiento se puede probar el resultado correspondiente a las densidades de transición.

Proposición 3.4.6 (Ecuación Backward de Kolmogorov II). *Sea X una solución de la ecuación (3.28). Suponga que la medida $P(s, y, t, dx)$ tiene una densidad $p(s, y, t, x)dx$, entonces*

$$\begin{aligned} \left(\frac{\partial p}{\partial s} + \mathcal{A}p \right) (s, y, t, B) &= 0, \quad (s, y) \in (0, t) \times \mathbb{R}^n, \\ P(s, y, t, x) &\rightarrow \delta_x, \quad \text{conforme } s \rightarrow t. \end{aligned}$$

La razón por la cual estas ecuaciones son llamadas ecuaciones *Backward* (regresivas), es que el operador diferencial está trabajando sobre las “variables regresivas” (s, y) .

Ahora se obtendrá una ecuación *forward* (progresiva) correspondiente, donde la acción del operador diferencial está sobre las “variables progresivas” (t, x) . Por simplicidad solo se considerará el caso escalar.

Suponga que X tiene una densidad de transición. Tome dos puntos fijos en el tiempo s y T con $s < T$. Considere una *función de prueba* arbitraria, es decir, una función diferenciable infinita $h(t, x)$ con soporte compacto en el conjunto $(s, T) \times \mathbb{R}$. De la fórmula de Itô se tiene

$$h(T, X_T) = h(s, X_s) + \int_s^T \left(\frac{\partial h}{\partial t} + \mathcal{A}h \right) (t, X_t) dt + \int_s^T \frac{\partial h}{\partial x} (t, X_t) dW_t.$$

Aplicando el operador esperado $E_{s,y}[\cdot]$, y usando el hecho que por el soporte compacto $h(T, x) = h(s, x) = 0$, se obtiene que

$$\int_{-\infty}^{\infty} \int_s^T p(s, y, t, x) \left(\frac{\partial h}{\partial t} + \mathcal{A} \right) h(t, x) dx dt = 0.$$

Integrando parcialmente con respecto a t y con respecto a x (para \mathcal{A}) resulta

$$\int_{-\infty}^{\infty} \int_s^T h(t, x) \left(-\frac{\partial h}{\partial t} + \mathcal{A}^* \right) p(s, y, t, x) dx dt = 0,$$

donde el operador adjunto \mathcal{A}^* es definido por

$$(\mathcal{A}^* f)(t, x) = -\frac{\partial}{\partial x} [\mu(t, x)f(t, x)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [\sigma^2(t, x)f(t, x)].$$

Puesto que esta ecuación se cumple para una *función de prueba* arbitraria, se tiene ya probada la siguiente proposición.

Proposición 3.4.7 (Ecuación Forward de Kolmogorov). *Suponga que la solución X de la ecuación (3.28) tiene una densidad de transición $p(s, y, t, x)$. Entonces p satisfará la ecuación forward de Kolmogorov*

$$\begin{aligned} \frac{\partial}{\partial t} p(s, y, t, x) &= \mathcal{A}^* p(s, y, t, x), \quad (t, x) \in (0, T) \times \mathbb{R}, \\ p(s, y, t, x) &\rightarrow \delta_y, \quad \text{conforme } t \rightarrow s. \end{aligned}$$

Esta ecuación es también conocida como la ecuación de Fokker-Planck. La versión multi-dimensional se obtiene fácilmente como

$$\frac{\partial}{\partial t} p(s, y, t, x) = \mathcal{A}^* p(s, y, t, x),$$

donde el operador adjunto \mathcal{A}^* está definido por

$$(\mathcal{A}^* f)(t, x) = - \sum_{i=1}^n \frac{\partial}{\partial x_i} [\mu_i(t, x) f(t, x)] + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} [C_{ij}(t, x) f(t, x)].$$

Se terminará esta sección con un ejemplo sencillo donde se utiliza la ecuación de Fokker-Planck.

Ejemplo. Considere un proceso de Wiener Estándar con coeficiente de difusión constante σ , es decir, la EDE

$$dX_t = \sigma dW_t.$$

La ecuación de Fokker-Planck para este proceso está dada por

$$\frac{\partial}{\partial t} p(s, y, t, x) = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}(s, y, t, x),$$

Es fácil ver que la solución está dada por la densidad Gaussiana

$$p(s, y, t, x) = \frac{1}{\sigma \sqrt{2\pi(t-s)}} \exp \left\{ -\frac{1}{2} \frac{(x-y)^2}{\sigma^2(t-s)} \right\}.$$

■

Capítulo 4

Dinámica de Portafolios: una aplicación financiera

En este capítulo se hará una descripción de un mercado financiero, el cual está conformado por diferentes activos. Se supondrá conocida la dinámica del precio de los activos con el objetivo principal de obtener la dinámica de un portafolio, que se conoce como portafolio autofinanciable. En tiempo continuo esta descripción resulta ser un poco delicada, por esta razón se comenzará el estudio del modelo en tiempo discreto, y, tomando la longitud del paso del tiempo tendiente a cero, se obtendrá el análogo en continuo.

4.1. Portafolios

Considere un mercado financiero donde el tiempo es dividido en períodos de longitud δ y donde la negociación solamente toma lugar sobre puntos en el tiempo discreto $n\Delta t$, $n = 0, 1, \dots$, y considere además un período fijo $[t, t + \Delta t)$. Este período donde $t = n\Delta t$ para algún n , será desde ahora referido como *período* t .

En adelante se hará la suposición adicional de que todos los activos son acciones, pero solamente por conveniencia lingüística.

A continuación se definirán algunos de los conceptos relevantes para la dinámica de portafolios:

- N : El número de diferentes tipos de activos.
- $h_i(t)$: Número de partes del tipo i que permanecen durante el período $[t, t + \Delta t)$.
- $h(t)$: El portafolio $h_1(t), \dots, h_N(t)$ que se mantiene durante el período t .
- $c(t)$: La cantidad de dinero gastado en el consumo por unidad de tiempo durante el período $[t, t + \Delta t)$.

- $S_i(t)$: El precio de una parte del tipo i durante el período de tiempo $[t, t + \Delta t)$.
- $V(t)$: El valor del portafolio h en el tiempo t .

La información y las decisiones en el modelo son estructuradas como sigue:

- En el tiempo t , es decir, al empezar el período, se tendrá un portafolio “viejo” $h(t - \Delta t) = \{h_i(t - \Delta t) : i = 1, \dots, N\}$ del período previo $t - \Delta t$.
- En el tiempo t se puede observar el vector de precios $S(t) = (S_1(t), \dots, S_N(t))$.
- En el tiempo t , después de tener observado $S(t)$, se elige un nuevo portafolio $h(t)$, para ser observado durante el período t . Al mismo tiempo se elige también la tasa de consumo $c(t)$ para el período t . Ambos, $h(t)$ y $c(t)$, se asumen como constantes en todo el período t .

Hasta el momento, sólo se han considerado activos que no pagan dividendos. En el caso en que los activos pagan dividendos las consideraciones son un poco más delicada, aunque se hará una descripción breve. En esta sección se consideraran únicamente los portafolios de consumo autofinanciables pares (h, c) , es decir, portafolios con infusión o retiro no exógena del dinero (aparte del término c). En otras palabras, la compra de un nuevo portafolio, así como todo el consumo, deberán ser financiadas solamente por la venta de los activos actuales del portafolio.

Suponga que se cuenta con un capital actual $V(t)$, es decir, el capital al comenzar el período t es igual al valor del portafolio anterior $h(t - \Delta t)$. Así se obtiene

$$V(t) = \sum_{i=1}^N h_i(t - \Delta t)S_i(t) = h(t - \Delta t)S(t) \quad (4.1)$$

La ecuación anterior simplemente establece que al principio del período t , el capital es igual al que se obtendría si se vendiera el portafolio “viejo” a precio de hoy. Se pueden usar los ingresos de esta venta para reinvertirlos en un nuevo portafolio $h(t)$. El costo del nuevo portafolio $h(t)$, el cual tiene que ser comparado con precios de hoy, está dado por

$$\sum_{i=1}^N h_i(t)S_i(t) = h(t)S(t),$$

donde el costo para la tasa de consumo $c(t)$ es dada por $c(t)\Delta t$. La ecuación de presupuesto para el periodo resulta ser

$$h(t + \Delta t)S(t) = h(t)S(t) + c(t)\Delta t \quad (4.2)$$

Se observa que la ecuación (4.2) puede expresarse como

$$S(t)\Delta h(t) + c(t)\Delta t = 0 \text{ donde } \Delta h(t) = h(t) - h(t - \Delta t) \quad (4.3)$$

Para obtener diferenciales de Itô se tiene que reformular la ecuación (4.3). Esto se hace sumando y restando el término $S(t - \Delta t)\Delta h(t)$ al lado izquierdo de la ecuación de presupuesto para obtener:

$$S(t - \Delta t)\Delta h(t) + \Delta S(t)\Delta h(t) + c(t)\Delta t = 0 \quad (4.4)$$

Ahora, se puede tomar $\Delta t \rightarrow 0$ en la ecuación (4.4), de lo que resulta

$$S(t)dh(t) + dh(t)dS(t) + c(t)dt = 0 \quad (4.5)$$

Si se toma $\Delta t \rightarrow 0$ en la ecuación (4.1) se obtiene

$$V(t) = h(t)S(t). \quad (4.6)$$

Al aplicar la fórmula de Itô sobre la ecuación (4.6) se sigue que

$$dV(t) = h(t)dS(t) + S(t)dh(t) + dS(t)dh(t) \quad (4.7)$$

Resumiendo, la ecuación (4.7) es la ecuación general para la dinámica de un portafolio arbitrario, y la ecuación (4.5) es la ecuación de presupuesto que se mantiene para todos los portafolios autofinanciables. Sustituyendo (4.5) en (4.7) se obtiene la expresión deseada, que se conoce como dinámica de un portafolio autofinanciable.

$$dV(t) = h(t)dS(t) - c(t)dt \quad (4.8)$$

En particular, si un portafolio no tiene tasa de consumo alguna, dicho portafolio tiene la siguiente V -dinámica.

$$dV(t) = h(t)dS(t)dt \quad (4.9)$$

A continuación se dará una definición más formal a los elementos centrales dados anteriormente.

Definición 4.1.1. Considere el proceso de precios N -dimensional $\{S(t) : t \geq 0\}$.

- I) Un portafolio estratégico o simplemente un portafolio, es cualquier proceso \mathfrak{F}_t^s -adaptado N -dimensional $\{h(t) : t \geq 0\}$.
- II) El portafolio se dice es *Markoviano* si es de la forma $h(t) = h(t, S(t))$ para alguna función $h : \mathbb{R}_+ \times \mathbb{R}^N \rightarrow \mathbb{R}^N$.

III) El proceso valor V^h correspondiente al portafolio h está dado por

$$V^h(t) = \sum_{i=1}^N h_i(t)S_i(t).$$

IV) Un proceso de consumo es cualquier proceso \mathfrak{F}_t^s -adaptado unidimensional $\{c(t) : t \geq 0\}$.

V) Un portafolio de consumo par (h, c) se dice *autofinanciable*, si el proceso valor V^h satisface la condición $dV^h(t) = \sum_{i=1}^N h_i(t)dS_i(t) - c(t)dt$ es decir, si $dV^h(t) = h(t)dS(t)dt - c(t)dt$.

En general, el portafolio $h(t)$ depende de las trayectorias de precio pasadas $\{S(u) : u \leq t\}$. En lo que sigue se estudiará casi exclusivamente portafolios Markovianos, es decir, aquellos portafolios para los cuales el valor sobre el tiempo t depende sólo de la fecha actual, y los valores actuales del vector de precios $S(t)$.

Para propósitos computacionales es conveniente describir un portafolio en términos relativos en lugar de en términos absolutos, como se hizo anteriormente. En otras palabras, en lugar de especificar el número absoluto de partes que se mantienen de cierta acción, se especificará la proporción relativa del valor total del portafolio que es invertido en la acción.

Definición 4.1.2. Para un portafolio dado h , el correspondiente portafolio relativo u es dado por

$$u_i(t) = \frac{h_i S_i(t)}{V^h(t)}, \text{ con } i = 1, 2, \dots, N \quad (4.10)$$

donde $\sum_{i=1}^N u_i(t) = 1$.

La condición para ser autofinanciable puede ser dada fácilmente en términos del portafolio relativo.

Lema 4.1.1. *Un portafolio de consumo par (h, c) es autofinanciable si y sólo si*

$$dV^h(t) = V^h(t) \sum_{i=1}^N u_i(t) \frac{dS_i(t)}{S_i(t)} - c(t)dt. \quad (4.11)$$

Lema 4.1.2. *Sea c un proceso de consumo, y suponga que existe un proceso escalar Z junto con un vector $q = (q_1, q_2, \dots, q_N)$ tal que*

$$dZ(t) = Z(t) \sum_{i=1}^N q_i(t) \frac{dS_i(t)}{S_i(t)} - c(t)dt \quad (4.12)$$

$$\sum_{i=1}^N q_i(t) = 1. \quad (4.13)$$

Ahora define un portafolio h por

$$h_i(t) = \frac{q_i(t)Z(t)}{S(t)}. \quad (4.14)$$

Entonces el proceso valor V^h está dado por $V^h = Z$, el par (h, c) es autofinanciable y el correspondiente portafolio relativo u es dado por $u = q$.

Demostración. Por definición el proceso V^h es dado por $V^h(t) = h(t)S(t)$, así las ecuaciones (4.13) y (4.14) resultan ser

$$V^h(t) = \sum_{i=1}^N h_i(t)S_i(t) = \sum_{i=1}^N q_i(t)Z(t) = Z(t) \sum_{i=1}^N q_i(t) = Z(t). \quad (4.15)$$

Sustituyendo (4.15) en (4.14) se observa que el portafolio relativo u correspondiente a h es dado por $u = q$. Usando este hecho, e insertando (4.15) en (4.12) se obtiene:

$$dV^h(t) = V^h(t) \sum_{i=1}^N h_i(t)dS_i(t) - c(t)dt,$$

el cual muestra que (h, c) es autofinanciable. ■

4.1.1. Dividendos

Considere los procesos $D_1(t), D_2(t), \dots$, donde $D_i(t)$ denota el dividendo acumulativo pagado por el tenedor por unidad de activo i durante el intervalo $(0, t]$. Si $D_i(t)$ tiene la estructura $dD_i(t) = \delta_i(t)dt$, para algún proceso δ_i , entonces se dice que el activo i paga un rendimiento de dividendo continuo.

El dividendo pagado por el tenedor por unidad de activo i durante $(s, t]$ está dado por $D_i(t) - D_i(s)$ y, para el caso de un rendimiento de dividendo se tiene que

$$D - i(t) = \int_0^t \delta_i(s)ds.$$

Se supone que todos los procesos dividendo tienen diferenciales estocásticas. Ahora se va a obtener la dinámica de un portafolio autofinanciable donde el proceso valor V está dado por $V(t) = h(t)S(t)$.

La diferencia entre este caso y el caso en el que no se paga dividendos es que la ecuación de presupuesto (4.2) ha sido modificada. Se debe tener en cuenta el hecho de que el dinero a disposición en el tiempo t consiste ahora de dos términos.

- El valor del portafolio “viejo”, como en el caso anterior, está dado por $h(t - \Delta t)S(t)$.

- Los dividendos ganados durante el intervalo $(t - \Delta t, t]$ están dados por

$$\sum_{i=1}^N h_i(t - \Delta t) [D_i(t) - D_i(t - \Delta t)] = h(t - \Delta t)\Delta D(t).$$

La nueva ecuación de presupuesto es dada por

$$h_i(t - \Delta t)S(t) + h_i(t - \Delta t)\Delta D(t) = h(t)S(t) - c(t)\Delta t. \quad (4.16)$$

Con los mismos argumentos mencionados con anterioridad, se obtiene la siguiente dinámica de portafolio autofinanciable

$$dV(t) = \sum_{i=1}^N h_i(t)dS_i(t) + \sum_{i=1}^N h_i(t)dD_i(t) - c(t)dt,$$

Definición 4.1.3. Este portafolio puede ser definido formalmente como sigue

- I) El proceso valor V^h está dado por

$$V^h(t) = \sum_{i=1}^N h_i(t)S_i(t). \quad (4.17)$$

- II) El proceso ganancia G está definido por

$$G(t) = S(t) + D(t). \quad (4.18)$$

- III) El portafolio de consumo par (h, c) se conoce como autofinanciable si

$$dV^h(t) = \sum_{i=1}^N h_i(t)dG_i(t) - c(t)dt. \quad (4.19)$$

De lo anterior se tiene el siguiente resultado

Lema 4.1.3. *En términos de los pesos del portafolio relativo, la dinámica de un portafolio autofinanciable puede ser expresado como*

$$dV^h(t) = V(t) \cdot \sum_{i=1}^N u_i(t) \frac{dG_i(t)}{S_i(t)} - c(t)dt. \quad (4.20)$$

4.2. Precios en condiciones de arbitraje

En esta sección se analizará un caso especial del modelo de mercado financiero que consiste de solo dos activos: Un activo libre de riesgo con proceso de precio B , y una acción con proceso de precio S . También se presentarán dos conceptos centrales en el estudio de los mercados financieros: los reclamos contingentes y el arbitraje.

Definición 4.2.1. El proceso de precio B es el precio de un activo libre de riesgo si tiene la dinámica

$$dB(t) = r(t)B(t)dt, \quad (4.21)$$

donde r es una función determinística.

La característica de un activo libre de riesgo es que no tiene el término de difusión dW .

También se puede escribir la dinámica de B como

$$\frac{dB(t)}{B(t)} = r(t)dt,$$

así el proceso B puede ser escrito en forma integral como

$$B(t) = B(0) \exp \left\{ \int_0^t r(s)ds \right\}.$$

Una interpretación natural para un activo libre de riesgo, es que este corresponda a un capital invertido en un banco con una tasa de interés r a corto plazo; un caso especial importante aparece cuando r es una constante determinística, en cuyo caso se puede interpretar B como el precio de un bono.

Suponga que el precio de una acción está dado por

$$dS(t) = S(t)\alpha(t, S(t))dt + S(t)\sigma(t, S(t))dW(t), \quad (4.22)$$

donde W es un proceso de Wiener y α, σ son funciones determinísticas dadas.

La función σ se conoce como la volatilidad de S , mientras que α es la tasa de retorno promedio de S .

Existe una clara diferencia entre el precio del activo riesgoso S descrito anteriormente y el precio del activo libre de riesgo. La tasa de retorno para B está dada por

$$\frac{dB(t)}{B(t) \cdot dt} = r(t),$$

este término es localmente determinístico en el sentido que, en el tiempo t , se tiene un conocimiento completo del retorno por simple observación de la tasa prevaleciente a corto plazo $r(t)$. Por otro lado la tasa de retorno para la acción S dada por

$$\frac{dS(t)}{S(t) \cdot dt} = \alpha(t, S(t)) + \sigma(t, S(t)) \frac{dW(t)}{dt},$$

no es observable en el tiempo t . Esta ecuación está compuesta por los términos $\alpha(t, S(t))$ y $\sigma(t, S(t))$, los cuales son ambos observables en el tiempo t , más el término de “ruido blanco” dW/dt , el cual es aleatorio. De manera opuesta al activo libre de riesgo, la acción tiene una tasa de retorno estocástica, incluso a escala infinitesimal.

El caso especial más importante de los modelos anteriores ocurre cuando r , α y σ son constantes determinísticas. Este es el famoso modelo de Black-Scholes que se estudiará mas detalladamente en el próximo capítulo.

Definición 4.2.2. El modelo de Black-Scholes está compuesto por dos activos con dinámicas dadas por

$$dB(t) = rB(t)dt, \quad (4.23)$$

$$dS(t) = \alpha S(t) + \sigma S(t)dW(t), \quad (4.24)$$

donde r, α y σ son constantes determinísticas.

4.2.1. Reclamos contingentes y arbitraje

Considere el modelo para un mercado financiero representado por las ecuaciones (4.21) y (4.22). Se hará una aproximación al problema principal que se estudiará en esta sección: “hallar el precio de derivados financieros”. Más adelante se dará una definición formal, pero primero se presentará uno de los derivados más simples, la *opción call Europea*.

Definición 4.2.3. Una *opción call Europea* con precio de ejercicio K y tiempo de maduración T sobre el activo subyacente S , es un contrato definido por las siguientes cláusulas:

- El tenedor de la opción tiene, durante el tiempo T , el derecho a comprar una parte de la acción subyacente con precio de ejercicio K al suscriptor de la opción.
- El tenedor de la opción no está en ninguna forma obligado a comprar la acción subyacente.
- El derecho de comprar la acción subyacente al precio K puede únicamente ser ejercido en el momento preciso T .
- Quien vende la opción, está obligado a vender la acción subyacente al precio K en el momento preciso T .

El precio de ejercicio K y el tiempo de maduración T son determinados en el momento de la emisión de la opción, que en este caso es habitualmente $t = 0$.

Una *opción put Europea* es una opción que da al tenedor el derecho a vender una parte del activo subyacente a un precio de ejercicio predeterminado. Para una *opción call americana* el derecho a comprar parte del activo subyacente puede ser ejercido en cualquier tiempo

antes del tiempo de maduración dado. El factor común de todos estos contratos es que están completamente definidos en términos del activo subyacente S , el cual hace natural llamarlos instrumentos derivados o reclamos contingentes. Ahora se dará la definición formal para un reclamo contingente.

Definición 4.2.4. Considere un mercado financiero con vector de precios S . Un *reclamo contingente* con fecha de maduración T , es una variable estocástica $X \in \mathfrak{F}_T^S$. Un reclamo contingente se llama un reclamo simple si este es de la forma $X = \Phi(S(t))$. La función Φ se llama función de contrato.

La opción call Europea es un reclamo contingente simple, cuya función contrato es dada por

$$\Phi(x) = \text{máx}\{x - K, 0\}.$$

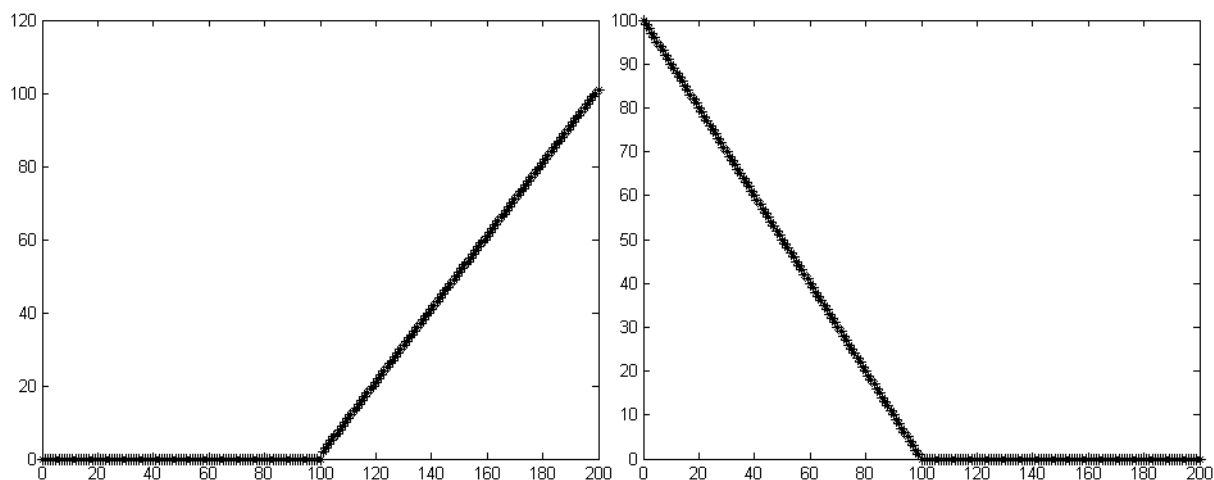


Figura 4.1: Con $K = 100$, función de contrato call Europea (a la izquierda) y función de contrato put Europea (a la derecha).

El valor exacto de la opción en el mercado dependerá por supuesto, del tiempo t y del precio $S(t)$ de la acción subyacente. El problema principal es determinar un precio “justo” del reclamo, para el cual se usará la notación estándar

$$\Pi(t; X) \tag{4.25}$$

para el proceso de precio del reclamo X . En ocasiones se suprimirá la X . En el caso de un reclamo simple algunas veces se escribirá $\Pi(t; \Phi)$.

Considere primero el caso particular para una call Europea, en un tiempo T fijo.

- i) Si $S(t) \geq K$, se puede tener un cierto beneficio ejerciendo la opción para comprar una parte de la acción subyacente. Esto costará K . Entonces inmediatamente se vende el activo sobre la acción cambiándolo al precio $S(T)$, dando así un beneficio neto de $S(T) - K$.

ii) Si $S(T) < K$ la opción no tendrá valor alguno (no se ejerce la compra).

Así se observa que el precio razonable $\Pi(T)$ para la opción es dado por

$$\Pi(T) = \text{máx}\{S(T) - K, 0\}. \quad (4.26)$$

Exactamente de la misma forma se observa que para un reclamo contingente X se tiene la relación $\Pi(T, X) = X$, y en el caso particular de un reclamo simple $\Pi(T, X) = \Phi(S(T))$.

Sin embargo, para cualquier tiempo $t < T$, es difícil saber cual es el precio correcto para un reclamo X . El precio de una opción, como el precio de cualquier otro activo, está determinado por el mercado, y debe ser mucho más complejo debido a las preferencias de vendedores y compradores respecto al riesgo de un mercado, y a las diferentes expectativas acerca de los precios futuros de la acción.

Definición 4.2.5. Una posibilidad de arbitraje en un mercado financiero es un portafolio autofinanciable tal que

$$\begin{aligned} V^h(0) &= 0 \\ V^h(T) &> 0. \end{aligned}$$

Se dice que el mercado es libre de arbitraje si no existen posibilidades de arbitraje. Una posibilidad de arbitraje es equivalente a la posibilidad de hacer una cantidad positiva de dinero de la nada con probabilidad 1.

Proposición 4.2.1. *Suponga que el proceso de precio $\Pi(t)$ es tal que no existen posibilidades de arbitraje en un mercado que consiste de $(B(t), S(t), \Pi(t))$, y que existe un portafolio autofinanciable h , tal que el proceso valor V^h tiene la dinámica*

$$dV^h(t) = k(t)V^h(t)dt, \quad (4.27)$$

donde k es un proceso adaptado. Entonces se deberá cumplir que $k(t) = r(t)$ para toda t , o de lo contrario existirá la posibilidad de arbitraje.

4.2.2. Justificación Semi-intuitiva

Se bosquejará el argumento y se supondrá por simplicidad que k y r son constantes donde $k > r$. Si se puede pedir prestado dinero del banco a una tasa r , este dinero es inmediatamente invertido en una estrategia de portafolio h , donde esta crecerá hasta la tasa k con $k > r$. Así la inversión neta en $t = 0$ es cero, mientras que la riqueza en cualquier tiempo $t > 0$ será positiva. En otras palabras se tiene la posibilidad de arbitraje. Si por otro lado $r > k$, se vende en corto el portafolio h y se invierte este dinero en el banco, y de nuevo esto da posibilidades de arbitraje. Para el caso en que k y r no sean constantes ni sean determinísticas se razona de la misma forma.

El punto principal de este razonamiento es que si un portafolio tiene un proceso valor cuya dinámica no contiene ningún proceso de Wiener, es decir, un portafolio libre de riesgo, entonces la tasa de retorno de este portafolio deberá ser igual a la tasa de interés a corto plazo. Para tomarlo de otra forma, la existencia de un portafolio h es para propósitos prácticos equivalente a la existencia de capital en un banco donde k es la tasa de interés a corto plazo. Se puede inferir del análisis anterior, que para un mercado libre de arbitraje puede existir solamente una tasa de interés a corto plazo.

Se tomará un ejemplo simple. Es obvio que para una opción call Europea se debe tener la relación $\Pi(t) \leq S(t)$ en un mercado libre de arbitraje, porque no se compraría una opción para comprar una parte del activo en una fecha posterior al precio K , si la parte del activo se puede comprar a menos precio que la opción. Para un argumento más formal, suponga que en algún tiempo t se tiene la relación $\Pi(t) > S(t)$, entonces simplemente se vende una opción. Una parte del dinero puede ser usada para comprar el activo subyacente y el resto invertirlo en el banco y esperar hasta el tiempo T . En esta forma se ha creado un portafolio autofinanciable con inversión neta cero en el tiempo t .

En el tiempo T se deberá el $\max\{S(T) - K, 0\}$ al tenedor de la opción, pero este dinero puede ser pagado vendiendo la acción. La riqueza neta en el tiempo T será $S(T) - \max\{S(T) - K, 0\}$, el cual es positivo, más el dinero invertido en el banco. De esta forma se tiene un arbitraje.

Para el caso de un reclamo contingente simple el argumento formal será presentado en el siguiente capítulo, pero se dará la idea general.

Para comenzar, se puede suponer que el precio $\Pi(t; X)$ en el tiempo t está determinado, de alguna manera, por las expectativas alrededor del precio futuro $S(T)$ de la acción. Como S es un proceso de Markov, tales expectativas están basadas en su retorno sobre el valor presente del proceso de precios (más que sobre la trayectoria completa en $[0, t]$).

De este modo se puede suponer que:

- I) El instrumento derivado en cuestión, puede ser comprado y vendido en un mercado.
- II) El mercado es libre de arbitraje.
- III) El proceso de precios para el activo derivado es de la forma

$$\Pi(t; X) = F(t, S(t)), \quad (4.28)$$

donde F es alguna función suave.

El objetivo es determinar cual función F se debería tomar para que el mercado compuesto por $S(t)$, $B(t)$ y $\Pi(t; X)$ sea libre de arbitraje.

Esquemáticamente se procede de la siguiente manera.

- 1) Se considera α, σ, Φ, F y r como dadas de manera exógena.

- II) Se describe la dinámica del valor del portafolio hipotético autofinanciable basado en los instrumentos derivados y la acción subyacente (no se invertirá nada, ni será prestado nada por el banco).
- III) Resulta que, se puede formar un portafolio autofinanciable cuyo proceso de valor tiene una diferencial estocástica que no contiene un proceso de Wiener. Esta será de la forma (4.27) anterior.
- IV) Como se ha supuesto ausencia de arbitraje se debe tener $k = r$.
- V) La condición $k = r$, tendrá la forma de una ecuación diferencial parcial con F como la función desconocida. Para que el mercado sea eficiente, F deberá ser solución de esta EDP.
- VI) La ecuación tiene una única solución, resultando la única fórmula de precio para el derivado, el cual es consistente con ausencia de arbitraje.

Capítulo 5

La Fórmula de Black-Scholes: otra aplicación financiera

El modelo central de la Teoría de Valoración de Opciones es el modelo de Black-Scholes (1973) en el que se muestra que sin realizar supuestos sobre las preferencias de los inversionistas, se puede obtener una expresión para el valor de las opciones que no depende directamente del rendimiento esperado de la acción subyacente ni de la opción. Esto se consigue mediante una dinámica perfecta en un mercado libre de arbitraje con suposiciones bastante fuertes. Las hipótesis sobre las que se sustenta el modelo de Black-Scholes configuran un escenario ideal en el que es posible la negociación continua, en unos mercados perfectos en los que el tipo de interés libre de riesgo es constante y el precio del activo subyacente se comporta como una variable aleatoria que sigue Proceso Browniano Geométrico. Desde la propuesta inicial del modelo de Black-Scholes han surgido numerosos modelos de valoración de opciones. Algunos de ellos plantean extensiones del modelo original, con el propósito de adecuarlo a la valoración de opciones distintas a las contempladas inicialmente, pero en su mayoría se trata de alternativas de valoración construidas sobre supuestos más cercanos a la realidad.

En este capítulo se presentará el modelo tradicional de Black-Scholes partiendo de un mercado compuesto por dos activos que permiten construir un portafolio autofinanciable, hasta plantear una Ecuación Diferencial Parcial para obtener una función suave F , que obligue a que el mercado compuesto por $[S(t), B(t), \Pi(t)]$ sea libre de posibilidades de arbitraje, resolviendo así el problema de valor de frontera utilizando la proposición de Feynman-Kac.

5.1. Ecuación de Black-Scholes

En esta sección se supondrá que el mercado dado *a priori* comprende dos activos con dinámicas dadas por:

$$dB(t) = rB(t); \quad (5.1)$$

$$dS(t) = S(t)\alpha(t, S(t)) dt + S(t)\sigma(t, S(t)) dW(t), \quad (5.2)$$

donde la tasa de interés a corto plazo r , es una constante determinística. Considere un reclamo contingente simple de la forma

$$X = \Phi(S(T)) \quad (5.3)$$

y suponga que este reclamo puede ser negociado en el mercado y que su proceso de precio $\Pi(t) = \Pi(t; \Phi)$ tiene la forma

$$\Pi(t) = F(t, S(t)) \quad (5.4)$$

para alguna función suave F .

El problema es descubrir cuál función F se debe considerar para que el mercado compuesto por $[S(t), B(t), \Pi(t)]$ sea libre de posibilidades de arbitraje.

Se comienza calculando la dinámica del precio del activo derivado, aplicando la fórmula de Itô a la ecuación (5.4) para obtener:

$$d\Pi(t) = F_t dt + F_s dS(t) + \frac{1}{2} \left[F_{ss} (dS(t))^2 \right] \quad (5.5)$$

Al sustituir (5.2) en (5.5) resulta

$$d\Pi(t) = F_t dt + F_s (S(t)\alpha(t, S(t))dt + S(t)\sigma(t, S(t))dW(t)) + \frac{1}{2} \left[F_{ss} (S(t)\sigma(t, S(t))dW(t))^2 \right]$$

Agrupando términos semejantes y multiplicando y dividiendo por $F(t, S(t))$ al lado derecho de la ecuación anterior, se sigue que

$$d\Pi(t) = F \frac{F_t + F_s S \alpha + \frac{1}{2} F_{ss} (S^2 \sigma^2) dt}{F} + F \frac{F_s S \sigma dW(t)}{F}$$

donde F , α , S y σ denotan respectivamente $F(t, S(t))$; $\alpha(t, S(t))$; $S(t)$ y $\sigma(t, S(t))$.

De este modo se obtiene la ecuación

$$d\Pi(t) = \alpha_{\Pi}(t)\Pi(t)dt + \sigma_{\Pi}(t)\Pi(t)dW(t) \quad (5.6)$$

donde los procesos $\alpha_{\Pi}(t)$ y $\sigma_{\Pi}(t)$ están definidos por:

$$\alpha_{\Pi}(t) = \frac{F_t + \alpha S F_s + \frac{1}{2} \sigma^2 S^2 F_{ss}}{F} \quad (5.7)$$

$$\sigma_{\Pi}(t) = \frac{\sigma S F_s}{F} \quad (5.8)$$

Aquí los subíndices denotan derivadas parciales y se ha usado la notación

$$\frac{\sigma S F_s}{F} = \frac{\sigma(t, S(t)) S(t) F_s(t, S(t))}{F(t, S(t))}$$

y así, de la misma forma, para los otros términos anteriores.

Ahora se formará un portafolio basado en dos activos: la acción subyacente y el activo derivado. Denotando el portafolio relativo por (u_s, u_{π}) y usando la ecuación (5.1) se obtiene la siguiente dinámica para el valor V del portafolio.

$$dV(t) = V\{u_s[\alpha dt + \sigma dW] + u_{\pi}[\alpha_{\pi} dt + \sigma_{\pi} dW]\} \quad (5.9)$$

donde se ha suprimido t .

Agrupando los factores de dt y dW se obtiene

$$dV = V[u_s \alpha + u_{\pi} \alpha_{\pi}] dt + V[u_s \sigma + u_{\pi} \sigma_{\pi}] dW \quad (5.10)$$

Se observa que ambos corchetes son lineales en los argumentos u_s y u_{π} y que además, la única restricción para el portafolio relativo es que $u_s + u_{\pi} = 1$, para toda t .

Se define el portafolio relativo por el sistema de ecuaciones lineales

$$u_s + u_{\pi} = 1 \quad (5.11)$$

$$u_s \sigma + u_{\pi} \sigma_{\pi} = 0 \quad (5.12)$$

Usando este portafolio se observa que por su misma definición, el término de tendencia dW en la V -dinámica de la ecuación (5.10) se desvanece completamente; dejando la ecuación

$$dV = V[u_s \alpha + u_{\pi} \alpha_{\pi}] dt \quad (5.13)$$

Así se ha obtenido un portafolio libre de riesgo, y debido a que se ha exigido que el mercado sea libre de arbitraje, se puede usar la proposición mencionada anteriormente para deducir la relación

$$u_s \alpha + u_{\pi} \alpha_{\pi} = r \quad (5.14)$$

La anterior es la condición de ausencia de arbitraje.

Es fácil ver que el sistema (5.11)-(5.12) tiene la solución

$$u_s = \frac{\sigma_\pi}{\sigma_\pi - \sigma} \quad (5.15)$$

$$u_\pi = \frac{-\sigma}{\sigma_\pi - \sigma} \quad (5.16)$$

Al utilizar la ecuación (5.8), resulta el portafolio explícitamente como

$$u_s(t) = \frac{S(t)F_s(t, S(t))}{S(t)F_s(t, S(t)) - F(t, S(t))} \quad (5.17)$$

$$u_\pi(t) = \frac{-F(t, S(t))}{S(t)F_s(t, S(t)) - F(t, S(t))} \quad (5.18)$$

Ahora se sustituye (5.7),(5.17) y (5.18) en la condición de ausencia de arbitraje (5.14). Después de algunos cálculos, se obtiene la ecuación

$$F_t(t, S(t)) + rS(t)F_s(t, S(t)) + \frac{1}{2}\sigma^2(t, S(t))S^2(t)F_{ss}(t, S(t)) - rF(t, S(t)) = 0$$

Además, se debe tener en cuenta la relación

$$\Pi(T) = \Phi(S(T))$$

Estas dos ecuaciones se tienen que mantener con probabilidad 1 para cada t fijo. Además se puede demostrar que bajo ciertas suposiciones, la distribución de $S(t)$ para cada $t > 0$ fijo tiene soporte en los enteros positivos en la recta real. Así $S(t)$ puede tomar cualquier valor, de modo que F tiene que satisfacer la siguiente EDP (determinística):

$$\begin{aligned} F_t(t, s) + rsF_s(t, s) + \frac{1}{2}s^2\sigma^2(t, s)F_{ss}(t, s) - rF(t, s) &= 0 \\ F(T, s) &= \Phi(s) \end{aligned}$$

De este modo F es una solución del problema de valor de frontera anterior y en consecuencia se tiene una aplicación directa de la proposición de Feynman-Kac II estudiada en el capítulo 3.

Recopilando los resultados anteriores, se ha probado la siguiente proposición, la cual es de hecho uno de los resultados más importantes estudiados en este capítulo.

Teorema 5.1.1 (Ecuación de Black-Scholes). *Suponga que el mercado es especificado por las ecuaciones (5.1) y (5.2), y que se quiere buscar el precio de un reclamo contingente de la forma (5.3). Entonces la única función de precios de la forma (5.4) la cual es consistente con la ausencia de arbitraje, se presenta cuando F es la solución de valor frontera e el dominio $[0, T] \times \mathbb{R}_+$.*

$$F_t(t, s) + rsF_s(t, s) + \frac{1}{2}s^2\sigma^2(t, s)F_{ss}(t, s) - rF(t, s) = 0 \quad (5.19)$$

$$F(T, s) = \Phi(s) \quad (5.20)$$

Es importante reafirmar el hecho de que se ha obtenido el precio X en la forma $\Pi(t, X) = F(t, S(t))$, es decir, el precio del reclamo está dado como una función del precio S del activo subyacente. Este razonamiento está relacionado con la idea básica explicada anteriormente, en la que el precio del activo derivado está determinado de una forma consistente con el precio del activo. No se presenta una fórmula de valoración absoluta para X . Por el contrario, se obtendrán precios que están alrededor del precio relativo, es decir, precios de los activos del derivado en términos del precio del activo subyacente. En particular, esto significa que para usar la técnica de precios con arbitraje se debe tener uno o varios procesos de precio en todos los subyacentes dados *a priori*.

Puede observarse un hecho importante sobre la ecuación de precios, y es que esta no contiene la tasa media de retorno local $\alpha(t, s)$ del activo subyacente. En particular, esto significa que cuando se valoran derivados, la tasa de retorno local del activo subyacente no desempeña un papel importante.

El único aspecto del proceso de precios del activo subyacente que es de alguna manera importante, es la volatilidad $\sigma(t, s)$. Así, para una volatilidad dada, el precio para un derivado fijo (como una opción call Europea) será exactamente el mismo sin importar si la acción subyacente tiene un 10 %, un 50 %, o incluso un -50 % de tasa de retorno.

5.2. Valoración de riesgo neutral

Considere un mercado dado por las ecuaciones

$$dB(t) = rB(t)dt \quad (5.21)$$

$$dS(t) = S(t)\alpha(t, S(t))dt + S(t)\sigma(t, S(t))d\bar{W}(t) \quad (5.22)$$

Y un reclamo contingente de la forma $X = \Phi S(T)$. Entonces, el precio libre de arbitraje está dado por $\Pi(t, \Phi) = F(t, S(t))$ donde F es la solución a la ecuación de precios (5.19)-(5.20).

Ahora se puede considerar de nuevo el problema de la ecuación de precios, y se observa que esta ecuación puede ser resuelta usando la fórmula de representación estocástica de Feynman-Kac.

Usando los resultados del capítulo 3, se observa que la solución dada por

$$F(t, s) = e^{-\gamma(T,t)} E[\Phi(X(t))] \quad (5.23)$$

donde el proceso X está definido por la dinámica

$$dX(u) = rX(u)du + X(u)\sigma(u, X(u))dW(u) \quad (5.24)$$

$$X(t) = s \quad (5.25)$$

donde W es un proceso de Wiener.

Es importante destacar que la EDE (5.24) es de la misma forma que la del proceso de precios S . El único cambio importante, es que mientras S tiene la tasa de retorno promedio α , el proceso X tiene la tasa de interés a corto plazo r como la tasa de retorno promedio.

El anterior proceso X es una herramienta técnica definida por el momento, y puede llamarse de cualquier manera porque aún no se ha contextualizado. En vista de la semejanza entre X y S , es posible hacer el cambio de variable S en lugar de X , mientras no se confunda el proceso S “real” de (5.22) con el “nuevo” proceso. Una forma de alcanzar esta meta es utilizar el siguiente procedimiento:

Se puede denotar la medida “objetiva” de la probabilidad que gobierna el modelo real (5.21)-(5.22) con la letra P . Así se puede decir que la P -dinámica para el proceso S es (5.22). Ahora se definirá otra medida de probabilidad Q bajo la cual el proceso S tiene una distribución de probabilidad diferente. Esta se constituye definiendo la Q -dinámica de S como

$$dS(t) = rS(t)dt + S(t)\sigma(t, S(t))dW(t) \quad (5.26)$$

donde W es un Q -proceso de Wiener. Para distinguir la medida bajo la cual se toma la esperanza, se introducen algunas notaciones convencionales.

Notación convencional

Se utilizarán las siguientes convenciones:

- Se identificará el operador esperanza, dejando E para denotar la esperanza tomada bajo la medida P ; donde E^Q denota el valor esperado tomado bajo la medida de probabilidad Q .
- Se identificará el proceso de Wiener. De este modo \bar{W} denotará un P -proceso de Wiener y W denotará un Q -proceso de Wiener.

La convención sobre W tiene la ventaja que es posible, a simple vista, decidir bajo cual medida P o Q , es dada alguna EDE. Se trabajará mucho más a menudo bajo la medida

Q que bajo la medida P y esta es la razón por la que el proceso de Q -Wiener, W tiene una notación más simple que el procesos de P -Wiener, \bar{W} . Usando esta notación, se puede establecer el siguiente resultado central para derivar precios.

Teorema 5.2.1. *El precio libre de arbitraje para el reclamo $\Phi(S(T))$ está dado por $\Pi(t, \Phi) = F(t, S(t))$, donde F está dada por la fórmula*

$$F(t, s) = e^{-\gamma(T-t)} E^Q[\Phi(S(T))] \quad (5.27)$$

donde la Q -dinámica para S está dada por la ecuación (5.26).

Una interpretación económica natural para la fórmula (5.27) es que el precio de un derivado, dada la fecha t de hoy y el precio s de la acción hoy, es calculado tomando el valor esperado del pago final $E_{t,s}^Q[\Phi(S(T))]$ y entonces descontarlo del valor esperado del valor presente usando el factor de descuento $e^{-\gamma(T-t)}$. Cuando se tome el valor esperado, no se debe hacer usando la medida de probabilidad objetiva P . En su lugar, se utilizará la Q -medida definida en (5.26). Esta Q -medida se conoce como “medida de riesgo ajustada” o también medida de martingala, y ésta será la terminología que se utilizará en adelante. La razón para el nombre es que bajo Q el proceso de descuento $\frac{S(t)}{B(t)}$ resulta ser una Q -martingala.

Proposición 5.2.2 (Valoración del riesgo neutral). *En el modelo de Black-Scholes, el proceso de precios $\Pi(t)$ para cada activo negociado, ya sea el subyacente o el activo derivado, tiene la propiedad de que el proceso de precios normalizado*

$$Z(t) = \frac{\Pi(t)}{B(t)}$$

es una martingala bajo la medida Q .

La fórmula (5.27) es algunas veces referida como la fórmula de valoración de riesgo neutral. Suponga que todos los agentes tienen riesgo neutral, entonces todos los activos demandarán una tasa de retorno igual a la tasa de interés a corto plazo, es decir, en un mundo de riesgo neutral el precio de la acción tendrá realmente la Q -dinámica anterior (más precisamente en este caso, se tendrá $Q = P$). Además, en un mundo de riesgo neutral el valor presente de un pago estocástico futuro, será igual al valor esperado de los pagos netos descontados del valor presente usando la tasa de interés a corto plazo.

5.3. La fórmula de Black-Scholes

En esta sección se especificará el modelo de la sección anterior para el caso del modelo de Black-Scholes

$$dB(t) = rB(t)dt \quad (5.28)$$

$$dS(t) = \alpha S(t)dt + \sigma S(t)d\bar{W}(t) \quad (5.29)$$

donde α y σ son constantes. De los resultados de las secciones previas se sabe que el precio libre de arbitraje para un reclamo contingente $\Phi(S(T))$ está dado por

$$F(t, s) = e^{-\gamma(T-t)} E_{t,s}^Q[\Phi(S(T))] \quad (5.30)$$

donde la Q -dinámica para S está dada por

$$dS(u) = rS(u)du + \sigma S(u)dW(u) \quad (5.31)$$

$$S(t) = s \quad (5.32)$$

En esta EDE se reconoce claramente el Movimiento Browniano Geométrico, por lo tanto se puede escribir $S(T)$ explícitamente como

$$S(T) = s \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma(W(T) - W(t))\right\} \quad (5.33)$$

Por tanto, $S(T)$ puede escribirse como $S(T) = s e^Y$ donde Y es una variable estocástica con una distribución

$$N\left[\left(r - \frac{1}{2}\sigma^2\right)(T-t), \sigma\sqrt{T-t}\right] \quad (5.34)$$

Y de este modo se obtiene la siguiente fórmula de precios

$$F(t, s) = e^{-\gamma(T-t)} \int_{-\infty}^{\infty} \Phi(s e^Y) f(y) dy \quad (5.35)$$

donde f es la ecuación de densidad para la variable estocástica Y anterior.

La fórmula (5.35) es una fórmula integral, que para una elección general de una función contrato Φ , deberá ser evaluada numéricamente.

Sin embargo, existen algunos casos particulares donde se puede evaluar (5.35) más o menos analíticamente, y el más conocido de estos es el caso de la opción call Europea, donde Φ tiene la forma $\Phi = \max[x - K, 0]$. En este caso es conveniente la normalización, de modo que $S(T)$ puede escribirse como

$$S(T) = s e^{\tilde{r}\tau + \sigma\sqrt{\tau}Z}$$

donde $\tilde{r} = \left(r + \frac{1}{2}\sigma^2\right)$, $\tau = T - t$ y Z es una variable normal estandarizada.

La integral en (5.35) resulta ser

$$\int_{-\infty}^{\infty} \max[s e^{\tilde{r}\tau + \sigma\sqrt{\tau}Z} - K, 0] \phi(z) dz$$

donde ϕ es la densidad de la distribución $N[0, 1]$, es decir,

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

El integrando de la integral anterior, se desvanece cuando

$$s e^{\tilde{r}\tau + \sigma\sqrt{\tau}Z} < K$$

es decir, cuando $z < z_0$, donde $z_0 = \frac{\ln\left(\frac{K}{s}\right) - \tilde{r}\tau}{\sigma\sqrt{\tau}}$.

La integral puede ser escrita

$$\int_{z_0}^{\infty} (s e^{\tilde{r}\tau + \sigma\sqrt{\tau}Z} - K)\phi(z)dz = A - B$$

donde

$$\begin{aligned} A &= \int_{z_0}^{\infty} s e^{\tilde{r}\tau + \sigma\sqrt{\tau}Z} \phi(z) dz \\ B &= \int_{z_0}^{\infty} K \phi(z) dz \end{aligned}$$

La integral B puede, obviamente, ser escrita como

$$K * P(Z \geq z_0)$$

Y utilizando la simetría para la distribución $N[0, 1]$, ésta puede ser reescrita como

$$K * P(Z \leq -z_0)$$

Se denotará como es común, la función de distribución acumulativa para la distribución $N[0, 1]$ por N , es decir,

$$N[x] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{z^2}{2}} dz$$

De este modo B puede escribirse como

$$B = K * N[-z_0]$$

Para la integral A , se completan los cuadrados en el exponente para así obtener

$$\begin{aligned} A &= \frac{s e^{\tilde{r}\tau}}{\sqrt{2\pi}} \int_{z_0}^{\infty} e^{\sigma\sqrt{\tau}z - \frac{1}{2}z^2} dz \\ &= \frac{s e^{\tilde{r}\tau}}{\sqrt{2\pi}} \int_{z_0}^{\infty} e^{-\frac{1}{2}(z - \sigma\sqrt{\tau})^2 + \frac{1}{2}\sigma^2\tau} dz \\ &= \frac{s e^{\tilde{r}\tau}}{\sqrt{2\pi}} \int_{z_0}^{\infty} e^{-\frac{1}{2}(z - \sigma\sqrt{\tau})^2} dz \end{aligned}$$

Aquí se reconoce la densidad para la distribución $N[\sigma\sqrt{\tau}, 1]$, de este modo se tiene que

$$A = s e^{r\tau} P(Z' \geq z_0)$$

Donde Z' sigue una distribución $N[\sigma\sqrt{\tau}, 1]$. Normalizando Z' para obtener una variable $N[0, 1]$ y de nuevo usando simetría, se puede escribir

$$A = s e^{r\tau} N[-z_0 - \sigma\sqrt{\tau}]$$

De este modo se ha calculado la integral en (5.35) y finalmente se tiene el siguiente resultado conocido como la **Fórmula de Black-Scholes**

Proposición 5.3.1 (Fórmula de Black-Scholes). *El precio de una opción call Europea con precio de ejercicio K y tiempo de maduración T está dado por la fórmula $\Pi(t) = F(t, S(t))$, donde*

$$F(t, s) = sN[d_1(t, s)] - e^{-r(T-t)}K * N[d_2(t, s)] \quad (5.36)$$

Además, N es la función de distribución acumulativa para la distribución $N[0, 1]$ y

$$d_1(t, s) = \frac{1}{\sigma\sqrt{T-t}} \left[\left(\frac{s}{K} \right) + \left(r + \frac{1}{2}\sigma^2 \right) (T-t) \right] \quad (5.37)$$

$$d_2(t, s) = d_1(t, s) - \sigma\sqrt{T-t} \quad (5.38)$$

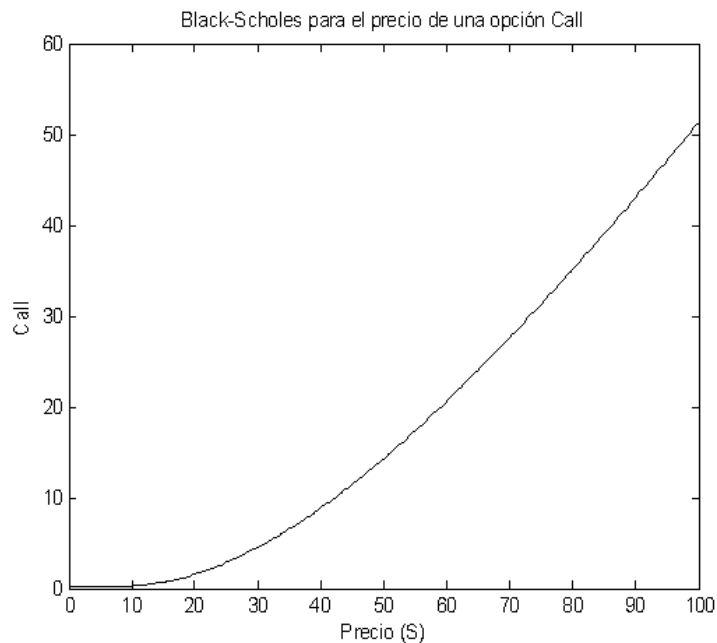


Figura 5.1: Black-Scholes para una opción call Europea

Capítulo 6

Procesos AR(1), GARCH(1,1) y de Reversión a la Media

Los rendimientos de las series financieras son prácticamente impredecibles, tienen una gran cantidad de valores extremos, y tanto los periodos de mayor agitación como los más pasivos están agrupados en el tiempo. Cuando la volatilidad es elevada, probablemente se mantenga elevada, y cuando es muy baja, también es probable que permanezca baja. Sin embargo, los periodos en los que se presentan estas fluctuaciones están limitados en el tiempo.

Estas características que se describen como impredecibilidad, exceso de curtosis y agrupamiento de la volatilidad, son precisamente las características para las cuales se diseñó el modelo ARCH propuesto por Engle (1982) el cual produce patrones dinámicos de reversion a la media que se pueden predecir y que además produce un mayor número de valores extremos de lo que se esperaría de una distribución normal estándar, pues los valores extremos durante el periodo de alta volatilidad son mayores de los que se hubiesen podido anticipar con un proceso de volatilidad constante.

Los modelos GARCH (ARCH generalizados) fueron propuestos por Bollerslev (1986) y se pueden emplear para describir la dinámica de volatilidad de casi todas las series de rendimientos financieros. Esto es cierto no sólo para las acciones de Estados Unidos, sino también para las acciones negociadas en la mayoría de los mercados desarrollados, para la mayoría de las acciones que se negocian en mercados emergentes y para la mayoría de los índices de los mercados de valores.

6.1. Modelos con Heterocedasticidad Condicional

Definición 6.1.1. Un proceso estocástico $\{Y_t; t \geq 0\}$ es estacionario de segundo orden si se cumplen las siguientes condiciones:

- I) $E[Y_t] = \mu$, para todo t donde μ es constante,
- II) $\text{var}[Y_t] = \sigma^2 < \infty \forall t$,
- III) $\text{cov}[Y_t, Y_{t-s}] = \gamma_s$, es decir, γ_s es una función definida positiva que sólo depende de la distancia entre t y $t - s$.

Definición 6.1.2. Se dice que un proceso $\{\varepsilon_t; t \geq 0\}$ es ruido blanco si se cumple:

- I) $E[\varepsilon_t] = 0$, para todo t ,
- II) $\text{var}[\varepsilon_t] = \sigma^2 < \infty \forall t$,
- III) $E[\varepsilon_t \cdot \varepsilon_{t'}] = 0, t \neq t'$.

6.1.1. El proceso AR(1)

Definición 6.1.3. Un proceso autoregresivo de orden uno, denotado AR(1), está dado por

$$Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t,$$

donde α, ρ son constantes y ε_t es un ruido blanco.

El proceso $Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$ también puede escribirse como

$$\begin{aligned} Y_t - \rho Y_{t-1} &= \alpha + \varepsilon_t \\ Y_t - \rho L Y_t &= \alpha + \varepsilon_t \\ (1 - \rho L) Y_t &= \alpha + \varepsilon_t \\ \Phi(L) Y_t &= \alpha + \varepsilon_t \end{aligned}$$

donde L denota el operador de rezago y $\Phi(L) = 1 - \rho L$ es el operador polinomial de rezagos.

Para que el proceso AR(1) sea estacionario debe cumplir que la raíz de $\Phi(L) = 0$ se encuentre por fuera del círculo unitario. Por lo que $1 - \rho L = 0 \Rightarrow L = 1/\rho$ tiene que tener a $|L| > 1 \rightarrow |1/\rho| > 1$, lo que implica que $|\rho| < 1$ garantiza la estacionariedad del proceso $Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$.

Teorema 6.1.1. Sea \mathfrak{F}_{t-1} la información generada por el proceso Y_t hasta el instante $t - 1$. Sea Y_t un proceso autoregresivo de orden uno, entonces,

- I) $E[Y_t] = \alpha/(1 - \rho)$ con $|\rho| < 1$,
- II) $\text{var}[Y_t] = \sigma_\varepsilon^2/(1 - \rho^2)$ con $|\rho| < 1$ donde $\sigma_\varepsilon^2 = E[\varepsilon_{t-i}^2]$ para toda i ,
- III) $E[Y_t | \mathfrak{F}_{t-1}] = \alpha + \rho Y_{t-1}$,

$$\text{IV) } \text{var}[Y_t | \mathfrak{F}_{t-1}] = \sigma_\varepsilon^2.$$

Observe que la razón $\text{var}[Y_t | \mathfrak{F}_{t-1}] / \text{var}[Y_t] < 1$ con $|\rho| < 1$, por lo cual la varianza condicional es menor que la no condicional aunque ambas son constantes en el tiempo. Teniendo en cuenta la observación, sería más apropiado considerar un proceso cuya varianza condicional no sea constante sino que dependa de la información disponible en cada instante, lo cual ayudaría a obtener estimadores más eficientes que presentarían menor amplitud en el rango de posibles valores para algún nivel de confianza dado.

Con estas consideraciones será más apropiado tener en cuenta otros procesos en los que se permita que la varianza condicional del proceso subyacente cambie a través del tiempo. La alternativa más favorable para este propósito es el proceso GARCH(1,1), el cual será descrito a continuación.

6.1.2. El proceso GARCH(1,1)

Definición 6.1.4. Considere el proceso $Y_t = \alpha + \beta Y_{t-1} + \varepsilon_t$ donde α, β son constantes. El proceso GARCH(1,1) para el ε_t está definido por $\varepsilon_t = h_t a_t$, donde $a_t \text{ iid} N(0, 1)$ y a_t, h_t son independientes, y $\varepsilon_t | \mathfrak{F}_{t-1} \sim N(0, h_t^2)$. Además $h_t^2 = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \theta_1 h_{t-1}^2$ con $\delta_0 > 0$ y $\delta_1, \theta_1 \geq 0$.

Teorema 6.1.2. Considere el proceso $Y_t = \alpha + \beta Y_{t-1} + \varepsilon_t$. Si ε_t sigue un proceso GARCH(1,1), entonces

$$\text{I) } E[\varepsilon_t^2] = \frac{\delta_0}{1 - (\delta_1 + \theta_1)}, \quad |\delta_1 + \theta_1| < 1,$$

$$\text{II) } E[h_t^2] = \frac{\delta_0}{1 - (\delta_1 + \theta_1)}, \quad |\delta_1 + \theta_1| < 1.$$

Demostración. II)

Teniendo ya el resultado de I), veamos que II) se cumple:

$$h_t^2 = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \theta_1 h_{t-1}^2; \text{ pero } h_{t-1}^2 = \delta_0 + \delta_1 \varepsilon_{t-2}^2 + \theta_1 h_{t-2}^2 \text{ y } h_{t-2}^2 = \delta_0 + \delta_1 \varepsilon_{t-3}^2 + \theta_1 h_{t-3}^2,$$

$$\begin{aligned} h_t^2 &= \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \theta_1 (\delta_0 + \delta_1 \varepsilon_{t-2}^2 + \theta_1 h_{t-2}^2) \\ &= \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \theta_1 \delta_0 + \theta_1 \delta_1 \varepsilon_{t-2}^2 + \theta_1^2 (\delta_0 + \delta_1 \varepsilon_{t-3}^2 + \theta_1 h_{t-3}^2) \\ &\vdots \\ &= \delta_0 (1 + \theta_1 + \theta_1^2 + \theta_1^3 + \dots) + \delta_1 (\varepsilon_{t-1}^2 + \theta_1 \varepsilon_{t-2}^2 + \theta_1^2 \varepsilon_{t-3}^2 + \dots) \\ &= \frac{\delta_0}{1 - \theta_1} + \delta_1 (\varepsilon_{t-1}^2 + \theta_1 \varepsilon_{t-2}^2 + \theta_1^2 \varepsilon_{t-3}^2 + \dots), \end{aligned}$$

donde $|\theta_1| < 1$ (para convergencia de la serie). Luego, si además $|\delta_1 + \theta_1| < 1$

$$\begin{aligned} E[h_t^2] &= \frac{\delta_0}{1 - \theta_1} + \delta_1 (E[\varepsilon_{t-1}^2] + \theta_1 E[\varepsilon_{t-2}^2] + \theta_1^2 E[\varepsilon_{t-3}^2] + \dots) \\ &= \frac{\delta_0}{1 - \theta_1} + \delta_1 (1 + \theta_1 + \theta_1^2 + \dots) \frac{\delta_0}{1 - (\delta_1 + \theta_1)} \\ &= \frac{\delta_0}{1 - \theta_1} + \frac{\delta_0}{1 - \theta_1} \cdot \frac{\delta_0}{1 - (\delta_1 + \theta_1)} = \frac{\delta_0}{1 - \theta_1} \left(\frac{1 - \theta_1}{1 - (\delta_1 + \theta_1)} \right) = \frac{\delta_0}{1 - (\delta_1 + \theta_1)}. \end{aligned}$$

■

6.2. Procesos de Reversión a la Media

Los procesos de Reversión a la Media son considerados como una elección natural para describir la dinámica de precios de algunos commodities tales como el petróleo, el aluminio, el gas natural y la energía eléctrica. Este tipo de procesos también han mostrado un buen desempeño en la descripción de tasas de interés y en el mercado de futuros.

En la literatura sobre economía y finanzas aparecen diferentes tipos de procesos de Reversión a la Media; entre los más conocidos se encuentran los modelos de uno, dos y tres factores. Para el propósito de este trabajo sólo se estudiará el modelo de un factor.

6.2.1. Procesos de Reversión a la Media de un solo factor

Considere la Ecuación Diferencial Estocástica Lineal Homogénea dada por

$$dS_t = \lambda(\mu - S_t)dt + \sigma S_t^r dW_t; \quad r = 0 \text{ o } 1,$$

donde $\lambda > 0$; $\sigma > 0$ y μ son constantes y W_t es un Movimiento Browniano Estándar Unidimensional.

Cuando $r = 0$, el ruido es aditivo, y en consecuencia la EDELH resulta ser $dS_t = \lambda(\mu - S_t)dt + \sigma dW_t$ que se conoce como proceso de Reversión a la Media del tipo Ornstein Uhlenbeck, que tiene por solución explícita el proceso

$$S_t = \mu + (S_0 - \mu)e^{-\lambda t} + \sigma e^{-\lambda t} \int_0^t e^{\lambda s} dW_s.$$

Puede determinarse la distribución invariante para S_t si se conoce $E[S_t]$ y $\text{var}[S_t]$.

$$\begin{aligned} E[S_t] &= \mu + (S_0 - \mu)e^{-\lambda t} + \sigma e^{-\lambda t} E \left[\int_0^t e^{-\lambda s} dW_s \right] \\ &= \mu + (S_0 - \mu)e^{-\lambda t}. \\ \text{var}[S_t] &= E \left[\left(\mu + (S_0 - \mu)e^{-\lambda t} + \sigma e^{-\lambda t} \int_0^t e^{\lambda s} dW_s \right)^2 \right] - \left(\mu + (S_0 - \mu)e^{-\lambda t} \right)^2 \\ &= 2\sigma e^{-\lambda t} (\mu + (S_0 - \mu)e^{-\lambda t}) \underbrace{E \left[\int_0^t e^{\lambda s} dW_s \right]}_{=0} + \sigma^2 e^{-2\lambda t} E \left[\int_0^t e^{\lambda s} dW_s \right]^2 \\ &= \sigma^2 e^{-2\lambda t} \int_0^t e^{2\lambda s} ds = \frac{\sigma^2}{2\lambda} (1 - e^{-2\lambda t}). \end{aligned}$$

Ahora $\mu + (S_0 - \mu)e^{-\lambda t} \rightarrow \mu$ conforme $t \rightarrow \infty$ y $\frac{\sigma^2}{2\lambda} (1 - e^{-2\lambda t}) \rightarrow \frac{\sigma^2}{2\lambda}$ conforme $t \rightarrow \infty$, en consecuencia S_t tiene una distribución invariante $N(\mu, \frac{\sigma^2}{2\lambda})$ para un S_0 arbitrario.

Cuando $r = 1$, el ruido es proporcional y en consecuencia la EDELH resulta ser

$$dS_t = \lambda(\mu - S_t)dt + \sigma S_t dW_t, \quad (6.1)$$

que se conoce simplemente como proceso de Reversión a la Media, donde μ es el nivel de equilibrio al largo plazo y λ es la velocidad de reversión.

Suponga que el valor esperado de S_t , denotado por \bar{S} , y que la Ecuación Diferencial Ordinaria dada por $d\bar{S} = \alpha(L - \bar{S})dt$. Entonces $\bar{S} = L$ es el nivel de equilibrio para esta Ecuación Diferencial Lineal Ordinaria, cuyas soluciones son dadas por $\bar{S}(t) = L + Ce^{-\alpha t}$, para alguna constante C . De este modo $\bar{S}(t) \rightarrow L$ conforme $t \rightarrow \infty$. En otras palabras, la media $\bar{S}(t)$ revierte al largo plazo a L ; de aquí el uso del término “Reversión a la Media”. En un horizonte de tiempo finito, L juega el papel de un atractor en el sentido de que, cuando $S_t > L$, el término de tendencia $\alpha(L - S_t) < 0$ y por lo tanto S_t decrece y, cuando $S_t < L$ un argumento similar establece que S_t crece.

La ecuación (6.1) tiene por solución explícita el proceso

$$S_t = S_0 \exp [-(\lambda + \sigma^2/2)t + \sigma W_t] + \lambda \mu \int_0^t \exp [-(\lambda + \sigma^2/2)(t - s) + \sigma(W_t - W_s)] ds$$

Ahora

$$\begin{aligned} E[S_t] &= S_0 E \left[\exp \left\{ -\left(\lambda + \frac{\sigma^2}{2}\right)t + \sigma W_t \right\} \right] + \lambda \mu E \left[\int_0^t \exp \left\{ -\left(\lambda + \frac{\sigma^2}{2}\right)(t - s) + \sigma(W_t - W_s) \right\} ds \right] \\ &= S_0 \exp \left\{ -\left(\lambda + \frac{\sigma^2}{2}\right)t + \frac{\sigma^2}{2}t \right\} + \lambda \mu \int_0^t \exp \left\{ -\left(\lambda + \frac{\sigma^2}{2}\right)(t - s) + \frac{\sigma^2}{2}(t - s) \right\} ds \\ &= S_0 e^{-\lambda t} + \lambda \mu \int_0^t e^{-\lambda(t-s)} ds \\ &= \mu + (S_0 - \mu)e^{-\lambda t}. \end{aligned}$$

Lo que indica que $\mu + (S_0 - \mu)e^{-\lambda t} \rightarrow \mu$ conforme $t \rightarrow \infty$.

Capítulo 7

Estimación de Modelos de Difusión: la conexión con el mundo “real”

La estimación de parámetros en los modelos de difusión es un punto clave para la aplicación de Ecuaciones Diferenciales Estocásticas a fenómenos asociados con el crecimiento poblacional, actividad neuronal, sismología y, particularmente, en economía y finanzas, cuyos intereses han permitido el desarrollo de este trabajo.

En este capítulo se hará una exposición de algunos métodos de estimación de modelos de difusión que son conocidos en la literatura, aunque no por ello son métodos apropiados o fáciles de implementar, pues en muchos de los casos no se obtienen estimadores eficientes o no se conoce con claridad la función de densidad de probabilidad de los procesos donde subyacen.

Se presentará el Método de Momentos y Generador Infinitesimal con sus características principales, y también los métodos basados en las funciones de escala y velocidad. Además se hará un breve recorrido por el Método de Momentos Simulados, el Método de Inferencia Indirecta y el Método de Máxima Verosimilitud.

Debido a su fácil implementación, familiaridad y características apropiadas de sus estimadores, se determinó utilizar el Método de Máxima Verosimilitud en el caso de aplicación que se presentará en el último capítulo, sin restar importancia a los métodos aquí estudiados.

7.1. Método de Máxima Verosimilitud

Las expresiones analíticas para las funciones de verosimilitud existen solamente para un número limitado de procesos de difusión. Debido a su maleabilidad, estos modelos han sido estudiados en la literatura, y a pesar de esto no proporcionan un muy buen ajuste a los precios observados de activos si no se consideran algunas características propias

de la heterocedasticidad condicional. De hecho, muchos procesos de difusión no explican características tales como volatilidad estocástica o el efecto de apalancamiento.

A estos procesos pertenece el Movimiento Browniano Geométrico, que aparece en la fórmula de Black-Scholes, el modelo de Ornstein-Uhlenbeck, usado por Vasicek para especificar la dinámica de las tasas de interés a corto plazo, y el modelo CIR.

7.1.1. Estimación de Máxima Verosimilitud para el Proceso de Ornstein-Uhlenbeck

Considere la Ecuación Diferencial Estocástica dada por

$$\begin{aligned} dy_t &= (\Phi - \lambda y_t)dt + \sigma dW_t \\ &= \lambda(\mu - y_t)dt + \sigma dW_t \end{aligned}$$

Donde W_t es un Movimiento Browniano Estándar. El proceso solución para esta ecuación tiene una contraparte en tiempo discreto dada por:

$$y_t = \mu [1 - \exp(-\lambda)] + \exp(-\lambda)y_{t-1} + \sigma \left(\frac{1 - \exp(-2\lambda)}{-2\lambda} \right)^{\frac{1}{2}} \epsilon_t$$

donde ϵ_t es un ruido blanco Gaussiano estandarizado. Esta ecuación corresponde a una representación Gaussiana AR(1) para el proceso (y_t, t) , $t \in \mathbb{Z}$.

7.1.2. Estimadores de Máxima Verosimilitud

Para aplicarse el método de Máxima Verosimilitud al modelo autoregresivo, primero se reparametrizará la representación autoregresiva como

$$y_t = \mu[1 - \rho] + \rho y_{t-1} + \eta \epsilon_t,$$

donde $\rho = \exp(-\lambda)$, $\mu = \frac{\phi}{\lambda}$ y $\eta = \sigma \left(\frac{-\exp(-\lambda)}{2\lambda} \right)^{\frac{1}{2}}$.

La función de log-verosimilitud para este proceso está dada por

$$\log(L) = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\eta^2) - \frac{1}{2\eta^2} \sum_{t=1}^T (y_t - \mu(1 - \rho) - \rho y_{t-1})^2$$

Al considerar $GRAD(\log(L)) = \left\langle \frac{\partial \log L}{\partial \eta}, \frac{\partial \log L}{\partial \rho}, \frac{\partial \log L}{\partial \mu} \right\rangle = \vec{0}$ se obtienen los estimadores de MV para los parámetros μ , ρ y η que son independientes asintóticamente y equivalentes

a

$$\begin{aligned}\hat{\mu}_T &= \frac{1}{T} \sum_{t=1}^T y_t = \bar{y}_T \\ \hat{\rho}_T &= \frac{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y}_T)(y_{t-1} - \bar{y}_T)}{\frac{1}{T} \sum_{t=1}^T (y_{t-1} - \bar{y}_T)^2} \\ \hat{\eta}_T^2 &= \frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_t^2\end{aligned}$$

donde los residuales son definidos por $\hat{\epsilon}_t = y_t - \bar{y}_T - \hat{\rho}_T(y_{t-1} - \bar{y}_T)$. Además, sus varianzas asintóticas están dadas por

$$\begin{aligned}V_{asy}(\hat{\mu}_T) &= \frac{\eta^2}{T(1 - \rho^2)} \\ V_{asy}(\hat{\rho}_T) &= \frac{1}{T}(1 - \rho^2) \\ V_{asy}(\hat{\eta}_T) &= \frac{2\eta^4}{T}\end{aligned}$$

De los estimadores MV de los parámetros μ , ρ y η , se infieren fácilmente los estimadores MV para los parámetros de interés

$$\begin{aligned}\hat{\lambda}_T &= -\log \hat{\rho}_T \\ \hat{\sigma}_T^2 &= -\frac{2 \log \hat{\rho}_T}{1 - \log \hat{\rho}_T^2} \hat{\eta}_T^2\end{aligned}$$

cuyas varianzas asintóticas son dadas por

$$\begin{aligned}V_{asy}(\hat{\lambda}_T) &= \left[\frac{\partial(-\log \rho)}{\partial \rho} \right]^2 V_{asy}(\hat{\rho}_T) \\ &= \frac{1}{T} \frac{1 - \rho^2}{\rho^2} \\ &= \frac{1}{T} \frac{1 - \exp(-2\lambda)}{\exp(-2\lambda)}\end{aligned}$$

Los estimadores $\hat{\lambda}_T$ y $\hat{\sigma}_T^2$ están asintóticamente correlacionados puesto que ambos dependen de $\hat{\rho}_T$.

7.2. Método de Momentos y Generador Infinitesimal

Cuando la función de verosimilitud no puede calcularse explícitamente, se pueden usar estimadores menos eficientes obtenidos mediante un criterio alternativo de optimización el cual admite una fórmula analítica. Un candidato natural para el criterio de optimización se presenta bajo las condiciones de momento asociadas con la ecuación diferencial estocástica. En esta sección, se revisarán varios métodos propuestos en la literatura en el contexto del proceso de difusión unidimensional. Observe que este marco implica un mercado completo y una fórmula única de valoración para los derivados, lo cual es incompatible por ejemplo, con la volatilidad estocástica.

7.2.1. Condiciones de Momento

Las condiciones de momento estudiadas por Hansen y Scheinkman (1995) están basadas en la siguiente proposición.

Proposición 7.2.1. *Considere un proceso y_t unidimensional que satisface la ecuación diferencial estocástica*

$$dy_t = \mu(y_t)dt + \sigma(y_t)dW_t$$

El generador infinitesimal A asociado con esta ecuación puede considerarse como un operador que transforma una función Φ de y en

$$\begin{aligned} A\Phi(y) &= \lim_{h \rightarrow 0} \frac{1}{h} E \left[\frac{\Phi(y_{t+h}) - \Phi(y_t)}{y_t = y} \right] \\ &= \frac{d\Phi(y)}{dy} \mu(y) + \frac{1}{2} \frac{d^2\Phi(y)}{dy^2} \sigma^2(y) \end{aligned}$$

Para un sistema grande de funciones $\Phi, \tilde{\Phi}$ se deben cumplir las siguientes condiciones de momento:

- I) $E[A\Phi(y_t)] = 0, \forall \Phi$
- II) $E \left[A\Phi(y_{t+1})\tilde{\Phi}(y_t) - \Phi(y_{t+1})A\tilde{\Phi}(y_t) \right] = 0, \forall \Phi, \tilde{\Phi}$

Por definición, el generador infinitesimal representa la tendencia infinitesimal del proceso transformado. La expresión diferencial para el generador sigue la fórmula de Itô aplicada a $\Phi(y_t)$.

Las condiciones de momento dadas en la proposición anterior conciernen a los momentos marginales de funciones no lineales de y_t (condición 1), y momentos cruzados para funciones no lineales de y_t y y_{t+1} (condición 2).

Como una ilustración, se efectuarán en detalle las condiciones de momento para las funciones exponenciales de y : $\Phi(y) = \exp(-ay)$, $\tilde{\Phi}(y) = \exp(-by)$.

Por la condición 1, se obtiene

$$E \left\{ \exp(-ay_t) \left[\mu(y_t) - \frac{a}{2} \sigma^2(y_t) \right] \right\} = 0, \quad \forall a$$

Por la condición 2 se obtiene

$$E [h^*] = 0, \quad \forall a, b$$

donde

$$h^* = a \exp(-ay_{t+1} - by_t) \left[\mu(y_{t+1}) - \frac{a}{2} \sigma^2(y_{t+1}) \right] - b \exp(-by_{t+1} - ay_t) \left[\mu(y_t) - \frac{a}{2} \sigma^2(y_t) \right]$$

7.2.2. Identificación

El método de momentos proporciona estimadores que pueden no ser completamente eficientes. Por lo tanto, es importante comprobar si el sistema anterior de condiciones de momento da suficiente información acerca de las funciones de tendencia μ y volatilidad σ . Sin embargo, pueden presentarse algunos problemas de identificación en la práctica. Por ejemplo al considerar la tendencia y la volatilidad parametrizadas como:

$$\mu(y; \theta) = \theta_0 \tilde{\mu}(y; \theta_1) \quad \sigma^2(y; \theta) = \theta_0 \tilde{\sigma}^2(y; \theta_1) \quad \text{con } \theta = \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$$

y aunque las condiciones 1 y 2 se pueden simplificar con respecto a θ_0 , estos parámetros no pueden ser identificados de las condiciones de momento dadas previamente.

Considerando todas las funciones Φ admisibles, se puede mostrar que la distribución marginal de y_t es identificable hasta un factor de escala. Todavía, el sistema de condiciones 2 no permite identificar la distribución conjunta de (y_t, y_{t+1}) hasta un factor escala.

Para ilustrar los problemas de identificación, se considera el siguiente modelo de difusión;

$$dy_t = (\alpha + \beta y_t) dt + \sigma y_t^\gamma dW_t$$

el cual fue propuesto por Chan (1992) para la tasa de interés a corto plazo.

Esta especificación abarca un número amplio de modelos de tiempo continuos (Broze, Scaillet y Zakoian 1995). Esta dinámica depende de cuatro parámetros: α , β , σ y γ .

La condición 1 para la función exponencial es:

$$E \left[\exp(-ay_t) \left(\alpha + \beta y_t - \frac{a}{2} \sigma^2 y_t^{2\gamma} \right) \right] = 0, \quad \forall a$$

Los parámetros que son identificables de la condición 1 son solamente γ , $\frac{\alpha}{\sigma^2}$ y $\frac{\beta}{\sigma^2}$.

7.2.3. Método de Momentos

El Método de Momentos es utilizado con frecuencia en procedimientos econométricos (Hansen 1982). Se presentará esta implementación para restricciones del tipo 1. Se seleccionarán n funciones a priori Φ_i para $i = 1, 2, \dots, n$.

Las condiciones de momento

$$E \left[\frac{d\Phi_i}{dy}(y_t)\mu(y_t; \theta) + \frac{1}{2} \frac{d^2\Phi_i}{dy^2}(y_t)\sigma^2(y_t; \theta) \right] = 0, \quad i = 1, 2, \dots, n$$

se cumplirán para el valor verdadero del parámetro. Se supone además, que algunas restricciones de identificación se imponen para eliminar el efecto del factor de escala (por ejemplo, $\theta_0 = 1$). Un estimador de momento es un valor del parámetro para el cual la condición de valor esperado anterior, “empareja” sus contrapartes empíricas.

Considere

$$\begin{aligned} \frac{d\Phi}{dy}(y_t) &= \left[\frac{d\Phi_1}{dy}(y_t), \dots, \frac{d\Phi_n}{dy}(y_t) \right]' \\ \frac{d^2\Phi}{dy^2} &= \left[\frac{d^2\Phi_1}{dy^2}(y_t), \dots, \frac{d^2\Phi_n}{dy^2}(y_t) \right]' \end{aligned}$$

Los estimadores de momento se derivan como la solución de la optimización

$$\hat{\theta}_T = \text{Arg} \min_{\theta} \sum_{t=1}^T (a \cdot b)$$

donde

$$\begin{aligned} a &= \left[\mu(y_t; \theta) \frac{d\Phi'}{dy}(y_t) + \frac{1}{2} \sigma^2(y_t; \theta) \frac{d^2\Phi'}{dy^2}(y_t) \right] \\ b &= \Omega \left[\mu(y_t; \theta) \frac{d\Phi}{dy}(y_t) + \frac{1}{2} \sigma^2(y_t; \theta) \frac{d^2\Phi}{dy^2}(y_t) \right] \end{aligned}$$

y Ω es una matriz definida positiva.

Cuando esta matriz es una identidad, la optimización produce:

$$\hat{\theta}_T = \arg \min_{\theta} \sum_{i=1}^n \left[\sum_{t=1}^T \mu(y_t) \frac{d\Phi_i}{dy}(y_t) \frac{1}{2} \sigma^2(y_t; \theta) \frac{d^2\Phi_i}{dy^2}(y_t) \right]^2$$

7.2.4. Descomposición Espectral del Generador Infinitesimal

Como se puede observar, el método anterior no permite la identificación de todos los parámetros de interés. Para evitar esta desventaja, se propone para analizar las características del generador infinitesimal, considerar la descomposición espectral (ver Demoura 1993, Hansen, Scheinkman y Touzi 1998, Florenz y Gourieroux 1998; Chen, Hansen y Scheinkman 1999). Más aún, esta aproximación proporciona estimadores no paramétricos de las funciones de tendencia y volatilidad.

Bajo las condiciones mínimas de regularidad, el operador infinitesimal admite una descomposición espectral, esto es, una sucesión de valores propios y funciones propias $\lambda_j, \Phi_j, j \geq 1$ tales que:

$$I) A\Phi_j(y) = \lambda_j\Phi_j, j \geq 1$$

$$II) \lambda_j, j \geq 1, \text{ son números reales positivos}$$

Ahora se organizan los valores propios. Se denota por λ_1 el valor propio más grande y por λ_2 el segundo más grande. Las condiciones que definen los pares $(\lambda_j, \Phi_j), j = 1, 2$ son:

$$\begin{cases} A\Phi_1(y) = \lambda_1\Phi_1(y) \\ A\Phi_2(y) = \lambda_2\Phi_2(y) \end{cases}$$

o equivalentemente

$$\begin{cases} \frac{d\Phi_1(y)}{dy}\mu(y) + \frac{1}{2}\frac{d^2\Phi_1(y)}{dy^2}\sigma^2(y) = \lambda_1\Phi_1(y) \\ \frac{d\Phi_2(y)}{dy}\mu(y) + \frac{1}{2}\frac{d^2\Phi_2(y)}{dy^2}\sigma^2(y) = \lambda_2\Phi_2(y) \end{cases}$$

Este es un sistema equivalente bivariado con respecto a la tendencia y a la volatilidad. De este modo, es equivalente estimar funciones μ y σ o estimar los dos primeros pares (λ_1, Φ_1) y (λ_2, Φ_2) no paramétricamente de las observaciones en tiempo discreto. La aproximación se basa en la interpretación de la predicción del generador infinitesimal

$$A\Phi(y) = \lim_{h \rightarrow 0} \frac{1}{h} E[\Phi(y_{t+h}) - \Phi(y_t) | y_t = y]$$

el cual representa la tendencia infinitesimal del proceso transformado. Después, se observa que la descomposición espectral de A está relacionada con la descomposición espectral del operador esperanza condicional T , el cual está asociado con la función Φ dada por:

$$T : \Phi \rightarrow T\Phi(y_t) = E[\Phi(y_{t+1}) | y_t]$$

Las funciones propias del operador esperanza condicional coinciden con las funciones propias del generador infinitesimal, mientras que sus valores propios están dados por $\tilde{\lambda}_i = \exp(\lambda_i)$

El sistema de ecuaciones anterior da la posibilidad de identificar no paramétricamente, las funciones de tendencia y volatilidad. Alternativamente, Ait-Sahalia (1996) propone usar la función de densidad marginal para identificar no paramétricamente la función volatilidad bajo el supuesto de una forma restrictiva de afinar la función de tendencia.

7.2.5. Descomposición Canónica no Lineal

La descomposición espectral de un operador esperanza está relacionado con la descomposición canónica no lineal de la fdp conjunta de (y_t, y_{t-1}) . Se denota por $f(y_t, y_{t-1})$ y $f(y_t)$ a las fdp conjunta y marginal respectivamente, si

$$\int \int \left[\frac{f(y_t, y_{t-1})}{f(y_t)f(y_{t-1})} \right]^2 f(y_t)f(y_{t-1}) dy_t dy_{t-1} < \infty$$

la fdp bivariada se puede descomponer como (Lancaster 1968)

$$f(y_t, y_{t-1}) = f(y_t)f(y_{t-1}) \left[1 + \sum_{j=1}^{\infty} \lambda_j \Phi_j(y_t) \Psi_j(y_{t-1}) \right]$$

donde las correlaciones canónicas λ_j , con j variando, son no negativas, y las direcciones canónicas Φ y Ψ , con j variando, satisfacen la restricción

$$\begin{aligned} E[\Phi_j(y_t)] &= E[\Psi_j(y_t)] = 0, \quad \forall j; & V[\Phi_j(y_t)] &= V[\Psi_j(y_t)] = 1, \quad \forall j \\ Cov[\Phi_j(y_t), \Phi_k(y_t)] &= 0, \quad \forall j \neq k; & Cov[\Psi_j(y_t), \Psi_k(y_t)] &= 0, \quad \forall j \neq k \end{aligned}$$

En particular, se deduce

$$f(y_t|y_{t-1}) = f(y_t) \left[1 + \sum_{j=0}^{\infty} \lambda_j \Phi_j(y_t) \Psi_j(y_{t-1}) \right]$$

Si y_t es un proceso de Markov de orden 1 se obtiene:

$$\begin{aligned} E[\Phi_j(y_t)|y_{t-1}] &= E[\Phi_j(y_t)] + \sum_{k=1}^{\infty} \lambda_k E[\Phi_j(y_t)\Phi_k(y_t)] \Psi_k(y_{t-1}) \\ &= \lambda_j \Psi_j(y_{t-1}) \end{aligned}$$

Los procesos de difusión unidimensionales son reversibles, esto es, sus propiedades distribucionales son idénticas en los tiempos ordinarios y reversos (Revuz y Yor 1990).

Esto implica que las direcciones canónicas ordinarias y las rezagadas son idénticas, es decir, $\Phi_j = \Psi_j, \forall j$. Así, la descomposición canónica no lineal resulta ser:

$$f(y_t, y_{t-1}) = f(y_t)f(y_{t-1}) \left[1 + \sum_{j=1}^{\infty} \lambda_j \Phi_j(y_t)\Phi_j(y_{t-1}) \right]$$

Por otra parte, el operador esperanza condicional es tal que

$$E[\Phi_j(y_t)|y_{t-1}] = \lambda_j \Phi_j(y_{t-1}), \forall j$$

Esto sugiere que los elementos espectrales del operador esperanza condicional pueden ser derivados de la descomposición canónica no lineal de la distribución bivariada.

7.2.6. Estimación para la Descomposición Espectral

En la literatura se proponen dos tipos de métodos de estimación no paramétrica. El Método del Filtro aproxima el operador esperanza condicional sobre un subespacio finito dimensional y requiere descomposición espectral para este operador aproximado. El segundo acercamiento descansa en la descomposición canónica no lineal de un estimador basado en el kernel de la fdp de $f(y_t, y_{t-1})$.

Método del Filtro

El primer paso de la aproximación consiste en proyectar las funciones Φ sobre un espacio dimensional (Chen, Hansen y Scheinkman 1999).

Se considera el espacio finito dimensional para las funciones

$$\Phi(y) = \sum_{k=1}^K b_k 1_{[a_k, a_{k+1}]}(y)$$

donde $[a_k, a_{k+1}]$ con $k = 1, 2, \dots, K$, es una partición dada de la recta real. Se puede aproximar el operador esperanza condicional proyectando estas funciones sobre el espacio generado por $1_{(a_k, a_{k+1})}(y_{t-1})$ con $k = 1, 2, \dots, K$.

Su contraparte empírica se obtiene fácilmente estimando por Mínimos Cuadrados Ordinarios el modelo

$$Z(t) = BZ(t-1) + u_t$$

$Z(t) = [1_{(a_1, a_2)}(y_t), \dots, 1_{(a_k, a_{k+1})}(y_t)]'$. Entonces, se realiza la descomposición espectral de la matriz de coeficientes estimada, \hat{B} , para encontrar los primeros valores y vectores propios asociados. Si $\hat{\varphi}_1, \dots, \hat{\varphi}_k$ son los vectores propios asociados, las correspondientes funciones propias son aproximadas por $\hat{\Phi}_k(y_t) = \hat{\Phi}'_k Z(t)$. (ver Chen, Hansen y Scheinkman 1999) para la escogencia de la partición y las propiedades asintóticas del estimador.

7.2.7. Estimación basada en el Kernel

Esta aproximación, fue introducida por Darolles, Florenz y Gourieroux (1998), estos autores proponen la aplicación del análisis canónico no lineal al estimador basado en el kernel para la fdp bivariada. La fdp estimada es

$$\hat{f}(y_t, y_{t-1}) = \frac{1}{Th^2} \sum_{\tau=1}^T K\left(\frac{y_\tau - y_t}{h}\right) K\left(\frac{y_\tau - y_{t-1}}{h}\right)$$

donde K es un kernel univariado y h es la longitud de la banda.

Se hace referencia a Darolles 1998 para una escogencia óptima de la longitud de banda h y las propiedades asintóticas de las correlaciones estimadas y las direcciones canónicas.

7.3. Métodos basados en las funciones de escala y la velocidad

Las aproximaciones basadas en las funciones de escala y velocidad consisten en reemplazar la parametrización inicial de una ecuación de difusión con funciones de tendencia y volatilidad por otra parametrización que sea más fácil para estimar utilizando los datos en tiempo discreto. Primero se definirán las funciones de escala y velocidad y se explicará su relevancia para la estimación de los modelos de difusión. Posteriormente, se mostrará cómo estimar estas funciones de las observaciones valoradas discretamente.

7.3.1. Escala intrínseca y tiempo discreto

El objetivo es simplificar la dinámica de un proceso de difusión aplicando escalas apropiadas y deformaciones del tiempo. La deformación de la escala cambia la unidad de medida en el espacio de todos los valores admisibles en el proceso, mientras que la deformación del tiempo modifica la unidad del tiempo de su trayectoria. Estas transformaciones, denotadas por S y M respectivamente, son aplicadas en el siguiente orden:

$$y_t \xrightarrow{S} z_t = S(y_t) \xrightarrow{M} x_t = z_{M(t)} = S[y_{M(t)}]$$

Proposición 7.3.1. *Existe un par de escala y de deformaciones del tiempo tales que $x_t = z_{M(t)} = W_t^*$ es un Movimiento Browniano. La deformación de escala puede ser fijada como*

$$S(y) = \int_0^y s(u) du \quad \text{donde} \quad s(y) = \exp - \int_0^y \frac{2\mu(u)}{\sigma^2(u)} du$$

La deformación del tiempo puede ser seleccionada para obtener

$$M(t) = \int_0^t m(y_u) du \quad \text{donde} \quad m(y) = \frac{1}{\sigma^2(y)s(y)}$$

Demostración

Primero, se estudiará el efecto de la deformación de la escala S .

Si $dy_t = \mu(y_t)dt + \sigma(y_t)dW_t$, se obtiene, por fórmula de Itô:

$$dz_t = \left[\mu(y_t)S'(y_t) + \frac{1}{2}\sigma^2(y_t)S''(y_t) \right] dt + \sigma(y_t)S'(y_t)dW_t$$

La transformación se establece para eliminar la tendencia:

$$\mu(y)S'(y) + \frac{1}{2}\sigma^2(y)S''(y) = 0$$

Se encuentra $S(y) = \int_0^y s(u)du$ donde $s(y) = \exp - \int_0^y \frac{2\mu(u)}{\sigma^2(u)}du$.

De esta manera, el proceso z_t satisface la ecuación de difusión

$$\begin{aligned} dz_t &= \sigma(y_t)S'(y_t)dW_t \\ &= \sigma[S^{-1}(z_t)]S'[S^{-1}(z_t)]dW_t \end{aligned}$$

Segundo, se introduce una deformación del tiempo del tipo

$$M(t) = \int_0^t m(y_u)du = \int_0^t \frac{m[S^{-1}(z_u)]}{S'[S^{-1}(z_u)]} dz$$

Se encuentra

$$\begin{aligned} dx_t &= \left| \frac{dM_t}{dz} \right|^{\frac{1}{2}} \sigma[S^{-1}(z_t)]S'[S^{-1}(z_t)]dW_t \\ &= \sigma(y_t)m(y_t)^{\frac{1}{2}}s(y_t)^{\frac{1}{2}}dW_t \end{aligned}$$

La deformación del tiempo se establece para transformar la volatilidad inicial en una volatilidad igual a 1:

$$\sigma(y)m(y)^{\frac{1}{2}}s(y)^{\frac{1}{2}} = 1 \iff m(y) = \frac{1}{\sigma^2(y)s(y)}$$

La proposición anterior muestra que es equivalente conocer las funciones de tendencia y volatilidad μ y σ^2 a las dos funciones S y M , llamadas función de escala y medida de la velocidad, respectivamente. Sin embargo estas funciones pueden ser difíciles de estimar. La siguiente proposición proporciona algunas interpretaciones útiles en términos de los momentos condicionales que pueden facilitar la aproximación.

Se define $T(a)$, como el primer tiempo en el cual el proceso alcanza el valor (a) , comenzando desde y en el tiempo $t = 0$. Considere dos constantes $a < b$ y un valor inicial $y \in [a, b]$.

Proposición 7.3.2.

$$I) u(y) = P[T(b) < T(a) | Y_0 = 0] = \frac{S(y) - S(a)}{S(b) - S(a)}, a < y < b$$

$$II) v(y) = E[\min[T(a), T(b)] | Y_0 = y]$$

$$= 2 \left\{ u(y) \int_y^b [S(b) - S(\xi)] m(\xi) d\xi + [1 - u(y)] \int_y^a [S(\xi) - S(a)] m(\xi) d\xi \right\}$$

Por otra parte las funciones u y v se pueden deducir del generador infinitesimal. Más precisamente ellos satisfacen las ecuaciones diferenciales dadas a continuación.

Proposición 7.3.3.

I) La función u es la solución de la ecuación diferencial

$$\mu \frac{du(y)}{dy} + \frac{1}{2} \sigma^2(y) \frac{d^2u(y)}{dy^2} = 0 \quad \text{para } a \leq y \leq b$$

con las condiciones de frontera $u(a) = u(b) = 1$

II) La función v es la solución de la ecuación diferencial

$$\mu \frac{dv(y)}{dy} + \frac{1}{2} \sigma^2(y) \frac{d^2v(y)}{dy^2} = -1 \quad \text{para } a \leq y \leq b$$

con las condiciones de frontera $v(a) = v(b) = 0$

7.3.2. Estimación de observaciones de valor discreto

Considere un proceso en tiempo continuo y_t , que satisface una ecuación de difusión, y suponga que se observan los datos en los cuales este proceso toma valores enteros, así como los valores asociados al proceso. Entonces, se pueden estimar fácilmente las expresiones

$$\begin{aligned} u(x) &= P[T(x+1) < T(x-1) | Y_0 = x] \\ v(x) &= E[\min[T(x+1), T(x-1)] | Y_0 = x] \end{aligned}$$

asociado don $a = x - 1$, $b = x + 1$, para algún entero x . De hecho, se denota por $\tau_j(x)$, $j = 1, 2, \dots, J_x$, a los datos para los cuales el proceso toma el valor de x , por $T_j(x+1)$ y $T_j(x-1)$, al primer dato después de $\tau_j(x)$ para los cuales el proceso toma los valores $(x+1)$ y $(x-1)$ respectivamente.

Los estimadores consistente son:

$$\begin{aligned} \hat{u}(x) &= \frac{1}{J_x} \sum_{j=1}^{J_x} 1_{T_j(x+1) < T_j(x-1)} \\ \hat{v}(x) &= \frac{1}{J_x} \sum_{j=1}^{J_x} [\min[T_j(x+1), T_j(x-1)] - \tau_j(x)] \end{aligned}$$

El sistema $\hat{u}(x)$, proporciona una aproximación a la función u para argumentos de valor entero. Se puede hacer una suavización para aproximar a u para argumentos reales. La misma aproximación se puede aplicar a la función v . Finalmente, la proposición anterior se puede usar para derivar no paramétricamente estimadores de las funciones de tendencia y volatilidad resolviendo el sistema bivariado

$$\begin{cases} \hat{\mu}(x) \frac{d\hat{u}(y)}{dy} + \frac{1}{2} \sigma^2 y \frac{d^2 \hat{u}(y)}{dy^2} = 0 \\ \hat{\mu}(x) \frac{d\hat{v}(y)}{dy} + \frac{1}{2} \sigma^2 y \frac{d^2 \hat{v}(y)}{dy^2} = -1 \end{cases}$$

para algún valor y .

Este procedimiento fue aplicado por Darolles, Gouriéroux, y Le Fol (2000) para datos de alta frecuencia sobre la acción de Elf-Aquitaine negociada en la Bolsa de Paris. Los autores consideraron contratos firmados, adheridos a órdenes de compra y encontraron que más del 92% de movimientos del precio eran más pequeños de una señal (la primera señal de los movimientos del precio es igual a $-1, 0, 1$). Los autores suponen que los datos de negociación valorados discretamente fueron obtenidos de un proceso de difusión subyacente por el valor del espacio de discretización.

De este modo, es necesario aplicar primero una función convexa de aumento al proceso y_t para conseguir un Movimiento Browniano bajo una deformación bien elegida del tiempo.

7.4. Método de Momentos Simulados

Contrario a la aproximación basada en el generador infinitesimal, el Método de Momentos Simulados, MMS, no requiere condiciones de momento en forma explícita (ver Duffie y Singleton 1993 para una aplicación particular). Se seleccionarán a priori algunas funciones $g_i(y_t, y_{t-1}, \dots, y_{t-p})$, para $i = 1, 2, \dots, n$ que dependan del presente y de p valores rezagados del proceso. Si este proceso satisface la ecuación diferencial estocástica

$$dy_t = \mu(y_t; \theta)dt + \sigma(y_t; \theta)dW_t$$

y_t es estacionario, entonces el valor esperado

$$E[g_i(y_t, y_{t-1}, \dots, y_{t-p})] = m_i(\theta) \quad i = 1, 2, \dots, n$$

es una función de parámetros. En general, las m_i tiene una forma complicada; así el estimador de momento estándar (Hansen 1982)

$$\hat{\theta}_T = \arg \min_{\theta} \sum_{t=1}^T [g_i(y_t, y_{t-1}, \dots, y_{t-p}) - m(\theta)]' \Omega \sum_{t=1}^T [g_i(y_t, y_{t-1}, \dots, y_{t-p}) - m(\theta)]$$

donde Ω es una matriz simétrica, definida positiva, que no puede ser calculada incluso numéricamente.

El principio del Método de Momentos Simulados consiste en reemplazar en el criterio pasado, la función desconocida m por una aproximación obtenida mediante simulaciones. Para este propósito, se requiere valores simulados $y_t^s(\theta)$ para $t = 1, 2, \dots, ST$, compatible con el modelo de difusión anterior y el valor θ para el parámetro. Una buena aproximación para $m(\theta)$ es

$$\hat{m}_T^s(\theta) = \frac{1}{ST} \sum_{t=1}^{ST} g [y_t^s(\theta), y_{t-1}^s(\theta), \dots, y_{t-p}^s(\theta)]$$

donde ST es suficientemente grande. Entonces, el estimador MMS es definido por

$$\hat{\theta}_T^s = \arg \min_{\theta} \sum_{t=1}^T [g(y_t, y_{t-1}, \dots, y_{t-p}) - \hat{m}_T^s(\theta)]' \Omega \sum_{t=1}^T [g(y_t, y_{t-1}, \dots, y_{t-p}) - \hat{m}_T^s(\theta)]$$

o equivalentemente por

$$\hat{\theta}_T^s = \arg \min_{\theta} [\hat{m}_T - \hat{m}_T^s(\theta)]' \Omega [\hat{m}_T - \hat{m}_T^s(\theta)]$$

donde $\hat{m} = \frac{1}{2} \sum_{t=1}^T g[y_t, y_{t-1}, \dots, y_{t-p}]$

Este método es esencialmente un acercamiento a la calibración que proporciona la mejor semilla posible entre los momentos empíricos, basado en las observaciones y éstas a su vez basadas en los valores simulados.

Esta aproximación requiere solamente la posibilidad de generar observaciones artificiales del modelo. De este modo, también puede ser usado en casos más generales, tales como modelos de difusión multivariados, modelos en tiempo continuo con volatilidad estocástica o modelos que incluyen componentes de salto.

7.5. Inferencia Indirecta

La inferencia indirecta también es un método basado en la simulación, aunque con una diferente función de criterio. Con argumentos similares a los dados anteriormente, este método también puede ser aplicado a los modelos con volatilidad estocástica y difusiones con saltos. Primero se presentará el principio general, seguido por una aplicación a un modelo en tiempo continuo univariado.

Suponga que se desea estimar el parámetros θ , de la función de log-verosimilitud

$$\sum_{t=1}^T \log f(y_t|y_{t-1}; \theta) = \sum_{t=1}^T \log f_t(\theta)$$

que no puede ser escrita en una forma explícita.

Empíricamente, una aproximación común consiste en reemplazar la función de verosimilitud por una expresión simple, por ejemplo:

$$\sum_{t=1}^T \log g(y_t|y_{t-1}; \theta) = \sum_{t=1}^T \log g_t(\theta)$$

y luego maximizar esta función modificada de log-verosimilitud.

Obviamente, el reemplazo de f por g induce un error de especificación y resultados probables en aproximaciones inconsistente para θ . La inferencia indirecta es una técnica basada en simulación diseñada para corregir el sesgo asintótico (ver Gouriéroux, Monfort, and Renault 1993; Galn and Tauchen 1996), la cual requiere dos entradas:

- i) Datos simulados compatibles con el modelo real y el valor θ para el parámetro.
- ii) Un modelo auxiliar (llamado un modelo instrumental)(Dhaene, Gouriéroux, and Scaillet 1998) con función de verosimilitud $\sum_{t=1}^T \log g_t(\beta)$ que es fácil de optimizar numéricamente.

La inferencia indirecta incluye los siguientes pasos:

- i) La estimación de parámetro auxiliar β del modelo instrumental y las observaciones:

$$\hat{\beta}_T = \arg \max_{\beta} \sum_{t=1}^T \log g_t(\beta)$$

- ii) La simulación del modelo verdadero para el valor θ del parámetro, produciendo una serie artificial de longitud ST : $y_t^s(\theta)$, para $s = 1, 2, \dots, ST$, y entonces la estimación del parámetro auxiliar del modelo instrumental y dato artificial es

$$\hat{\beta}_{ST}^s(\theta) = \arg \max_{\beta} \sum_{t=1}^{ST} \log g[y_t^s(\theta)|y_{t-1}^s(\theta); \beta]$$

- III) El último paso es el paso de calibración, construyendo por inferencia indirecta el estimador para θ definido por:

$$\hat{\theta}_T^s(\theta) = \arg \min_{\theta} \left[\hat{\beta}_T^s - \hat{\beta}_{sT}^s(\theta) \right]' \Omega \left[\hat{\beta}_T^s - \hat{\beta}_T^s(\theta) \right]$$

donde Ω es una matriz de pesos.

7.6. Método de los Momentos Eficientes

Una aproximación alternativa fue introducida por Gallant y Tauchen (1996). Está basada en un modelo auxiliar $g(y_t|y_{t-1}; \beta)$ y hace uso del estimador de quiasimáxima verosimilitud (QMV) $\hat{\beta}_T$. Este estimador satisface la condición de primer orden

$$\frac{1}{T} \sum_{t=1}^T \frac{\partial \log g_t(\hat{\beta}_T)}{\partial \beta} = 0$$

La cual converge asintóticamente a la solución β_0^* para la condición de momento límite

$$E \left[\frac{\partial \log (y_t|y_{t-1}; \beta_0^*)}{\partial \beta} \right] = 0$$

que se conoce como “valor pseudo-verdadero”. La condición de momento para la ecuación anterior se asemeja a la condición del método de momentos simulados. Más precisamente, Gallan y Tauchen consideraron las series $y_t^s(\theta)$, para $s = 1, 2, \dots, ST$, generado bajo el modelo verdadero y encontraron la solución de la siguiente optimización:

$$\hat{\theta}^s = \arg \min_{\theta} \left[\frac{1}{ST} \sum_{t=1}^T \sum_{s=1}^S \frac{\partial \log g(y_t^s(\theta)|y_{t-1}^s; \hat{\beta}_T)}{\partial \beta} \right] \Omega \left[\frac{1}{ST} \sum_{t=1}^T \sum_{s=1}^S \frac{\partial \log g(y_t^s(\theta)|y_{t-1}^s; \hat{\beta}_T)}{\partial \beta} \right]$$

donde Ω es una matriz de pesos. Se ha probado que los estimadores $\hat{\theta}^s$ y $\hat{\theta}^s$ dados anteriormente son asintóticamente equivalentes. Esta es una consecuencia de la aproximación de la función objetivo como el test estadístico de Wald y un test estadístico de multiplicadores de Lagrange. La eficiencia del estimador se incrementa con el número de momentos considerados. En el caso límite para un modelo auxiliar con un número infinito de parámetros, se alcanza eficiencia completa. Esto explica el término Método de los Momentos Eficientes (MME) usado como una referencia en la literatura (Gallant y Tauchen 1996).

7.6.1. Aplicación a modelos de tiempo continuo

Considere un proceso en tiempo continuo que satisface en la ecuación diferencial estocástica

$$dy_t = \mu(y_t; \theta)dt + \sigma(y_t; \theta)dW_t$$

donde W_t es un Movimiento Browniano Estándar.

Cuando las observaciones disponibles son muestreadas para datos enteros $1, 2, \dots, T$, es generalmente imposible determinar la forma analítica de una función de verosimilitud.

Una aproximación común consiste en reemplazar el modelo en tiempo continuo inicial por su discretización de Euler

$$y_t = y_{y-1} + \mu(y_{t-1}; \beta) + \sigma(y_{t-1} : \beta)\epsilon_t$$

donde ϵ_t es un ruido blanco Gaussiano. A continuación, se estimará β por el método de la máxima verosimilitud estudiado anteriormente y probablemente se concluya que el parámetro estimado $\hat{\beta}$ proporciona una buena aproximación del θ desconocido. Sin embargo, puesto que la discretización de Euler es una aproximación al modelo anterior, éste es inespecífico causando un sesgo asintótico para este estimador, el cual puede ser arbitrariamente grande. Sin embargo, la discretización de Euler puede ser usada como un modelo instrumental en el método de inferencia indirecta, la idea es introducir una segunda discretización de Euler que involucra a cada unidad corta de tiempo, por ejemplo $\delta = \frac{1}{10}$. Más precisamente, se define el proceso $y_t^\delta, t = k\delta$, k variando tal que

$$y_{(k+1)\delta}^\delta = y_{k\delta}^\delta + \delta\mu(y_{k\delta}^\delta, \theta) + \sigma(y_{k\delta}^\delta, \theta)\sqrt{\delta}\epsilon_k^\delta$$

donde ϵ_k^δ, k variando, es un ruido blanco Gaussiano. Esta discretización de Euler más fina puede ser usada para simular los procesos en tiempo continuo. Se denotará por $y_{k\delta}^{\delta,S}(\theta)$, con $k = 1, 2, \dots, \frac{T}{\delta}$, a la trayectoria simulada correspondiente al valor del parámetro θ y $y_t^s(\theta), t = y_t^{\delta,S}(\theta), t, t = 1, 2, \dots, T$ a los valores correspondientes de los datos observados.

La aproximación por Inferencia Indirecta, incluye los tres pasos de estimación listados a continuación:

I) Estimación para β de las observaciones

$$\hat{\beta}_T = \text{Arg mín}_{\beta} \sum_{t=1}^T \left\{ \frac{-1}{2} \log \sigma^2(y_{t-1}, \beta) - \frac{-1}{2} \frac{[y_t - \mu(y_{t-1}, \beta)]^2}{\sigma^2(y_{t-1}, \beta)} \right\}$$

II) Estimación para β de las simulaciones

$$\hat{\beta}_T^s(\theta) = \arg \min_{\beta} \sum_{t=1}^T \left\{ \frac{-1}{2} \log \sigma^2(y_{t-1}^s(\theta), \beta) - \frac{-1}{2} \frac{[y_t^s(\theta) - \mu(y_{t-1}^s(\theta), \beta)]^2}{\sigma^2(y_{t-1}^s(\theta), \beta)} \right\}$$

III) Calibración $\hat{\theta}_T^s(\theta) = \text{Arg mín}_{\beta} \left[\hat{\beta}_T - \hat{\beta}_T^s(\theta) \right]' \Omega \left[\hat{\beta}_T - \hat{\beta}_T^s(\theta) \right]$ donde Ω es una matriz de pesos dada.

Capítulo 8

Aplicación: el caso del aluminio

Como se expuso anteriormente, y luego de estudiar algunos métodos de estimación de modelos de difusión, se presentara un caso de aplicación en el cual se implementará el método de máxima verosimilitud para estimar los parámetros de la Ecuación Diferencial Estocástica asociada al proceso de reversión a la media del tipo Ornstein Uhlenbeck.

Considere la Ecuación Diferencial Estocástica dada por:

$$dy_t = (\phi - \lambda y_t)dt + dW_t \quad (8.1)$$

Donde $\phi, \lambda > 0$ y $k = \sqrt{2\lambda\theta}$ son constantes y W_t es un Movimiento Browniano Estándar Unidimensional. Suponga que y_t sigue un proceso con especificación GARCH(1,1), es decir

$$y_t = \alpha + \beta y_{t-1} + \varepsilon_t,$$

donde $\varepsilon_t = a_t h_t$ y $a_t \sim iidN(0, 1)$. Además $h_t^2 = \delta_0 + \delta_1 \varepsilon_{t-1}^2 + \theta_1 h_{t-1}^2$ con $\delta_0 > 0$, $\delta_1, \theta_1 \geq 0$.

Para resolver ésta Ecuación Diferencial Estocástica Lineal No Homogénea, es muy útil considerar primero una ecuación determinística asociada sin ruido, esto es, bajo la suposición $k = 0$.

La ecuación $dy_t = (\phi - \lambda y_t)dt$, es una Ecuación Diferencial Ordinaria Lineal en y_t . De esta manera, dicha ecuación admite una solución general de la forma

$$y_t = ce^{-\lambda t} + \frac{\phi}{\lambda},$$

donde c es una constante arbitraria.

Para resolver la Ecuación Diferencial Estocástica inicial, se introduce un cambio de variable como sigue:

$$y_t = \xi_t e^{-\lambda t} + \frac{\phi}{\lambda}.$$

Al utilizar el lema de Itô se obtiene

$$\begin{aligned} dy_t &= d\xi_t e^{-\lambda t} - \lambda \xi_t e^{-\lambda t} dt \\ &= d\xi_t e^{-\lambda t} - \lambda \left(y_t - \frac{\phi}{\lambda} \right) dt. \end{aligned}$$

Además

$$d\xi_t e^{-\lambda t} + (\phi - \lambda y_t) dt = (\phi - \lambda y_t) dt = (\phi - \lambda y_t) dt + k dW_t,$$

en consecuencia $d\xi_t = k \exp(\lambda t) dW_t$.

De esta forma, para cualquier par de datos $t < t'$, se puede expresar ξ_t como una función de $\xi_{t'}$ y de las realizaciones del ruido entre t' y t :

$$\xi_t = \xi_{t'} + \int_{t'}^t k e^{\lambda u} dW_u, \quad t' < t.$$

Como $y_t = \xi_t e^{-\lambda t} + \frac{\phi}{\lambda}$, se obtiene

$$y_t = \xi_{t'} e^{-\lambda t} + \int_{t'}^t k e^{\lambda u} dW_u e^{-\lambda t} + \frac{\phi}{\lambda}$$

pero $\xi_t = \left(y_{t'} - \frac{\phi}{\lambda} \right) e^{\lambda t'}$ entonces

$$\begin{aligned} y_t &= \left(y_{t'} - \frac{\phi}{\lambda} \right) e^{\lambda t'} e^{-\lambda t} + \int_{t'}^t k e^{\lambda u} dW_u e^{-\lambda t} + \frac{\phi}{\lambda} \\ &= \left[y_{t'} e^{\lambda t'} - \frac{\phi}{\lambda} e^{\lambda t'} \right] e^{-\lambda t} + \frac{\phi}{\lambda} + k e^{-\lambda t} \int_{t'}^t e^{\lambda u} dW_u \\ &= y_{t'} e^{-\lambda(t-t')} + \frac{\phi}{\lambda} \left[1 - e^{-\lambda(t-t')} \right] + k e^{-\lambda t} \int_{t'}^t e^{\lambda u} dW_u. \end{aligned}$$

Proposición 8.0.1. *Las soluciones de la Ecuación Diferencial Estocástica $dy_t = (\phi - \lambda y_t) dt + k dW_t$ para cualquier valor $t < t'$, están dadas por el proceso solución*

$$y_t = e^{-\lambda(t-t')} y_{t'} + \frac{\phi}{\lambda} \left[1 - e^{-\lambda(t-t')} \right] + k \int_{t'}^t e^{-\lambda(t-u)} dW_u.$$

Corolario 8.0.2. *Para un proceso del tipo Ornstein Uhlenbeck, $y_t - e^{-\lambda(t-t')} y_{t'}$ es independiente de $y_{t'}$.*

Ahora suponga que se quiere describir la relación anterior para el caso en que $t' = t - 1$. De esta manera se obtiene

$$e^{-\lambda} y_{t-1} + \frac{\phi}{\lambda} \left[1 - e^{-\lambda} \right] + k \int_{t-1}^t e^{-\lambda(t-u)} dW_u.$$

Observe que la integral $I_t = k \int_{t'}^t e^{-\lambda(t-u)} dW_u$, con t variando, es gaussiana e independiente con media y varianza dadas como sigue:

$$\begin{aligned} E[I_t] &= E \left[k \int_{t'}^t e^{-\lambda(t-u)} dW_u \right] = 0; \\ \text{var}[I_t] &= \text{var} \left[k \int_{t'}^t e^{-\lambda(t-u)} dW_u \right] = E \left[\left(k \int_{t'}^t e^{-\lambda(t-u)} dW_u \right)^2 \right] \\ &= k^2 \int_{t'}^t e^{-2\lambda(t-u)} du = k^2 \left[\frac{e^{-2\lambda(t-u)}}{-2\lambda} \right]_{t'}^t \\ &= \frac{k^2}{2\lambda} \left(e^{-2\lambda(t-t)} - e^{-2\lambda(t-t')} \right) = \frac{k^2}{2\lambda} (1 - e^{-2\lambda}). \end{aligned}$$

Se sigue que $I_t \sim N \left(0, \frac{k^2}{2\lambda} (1 - \exp(-\lambda)) \right) = k \left(\frac{1 - \exp(-2\lambda)}{2\lambda} \right) N(0, 1)$.

Luego se puede escribir y_t como:

$$y_t = e^{-\lambda} y_{t-1} + \frac{\phi}{\lambda} [1 - e^{-\lambda}] + (\theta(1 - e^{-2\lambda}))^{\frac{1}{2}} \varepsilon_t,$$

donde $(\varepsilon_t, t \in \mathbb{Z})$ es un ruido blanco Gaussiano con varianza 1.

Corolario 8.0.3. Para $\lambda > 0$, el proceso en tiempo discreto $(y_t, t \in \mathbb{Z})$ es un proceso auto-regresivo Gaussiano de orden 1, con media $\frac{\phi}{\lambda}$, coeficiente auto-regresivo $\exp(-\lambda)$ y varianza $\frac{\sigma^2}{2\lambda}(1 - \exp(-2\lambda))$.

Este resultado puede ser extendido para probar que, para $\lambda > 0$, el proceso en tiempo continuo $(y_t, t \in \mathbb{R})$ es también estacionario y Gaussiano. De este modo el proceso OU es el análogo del tiempo continuo para un procesos auto regresivo Gaussiano AR(1).

8.1. Estimadores de Máxima Verosimilitud

Para aplicar el método de Máxima Verosimilitud estudiado en el capítulo anterior al modelo AR(1), primero se reparametriza la representación auto-regresiva como

$$y_t = \mu[1 - \rho] + \rho y_{t-1} + \eta \varepsilon_t. \quad (8.2)$$

donde $\rho = \exp(-\lambda)$, $\mu = \frac{\phi}{\lambda}$, $\eta = (\theta(1 - \rho^2))^{\frac{1}{2}}$.

La ecuación (8.2) puede escribirse como

$$x_t = y_t - \mu[1 - \rho] - \rho y_{t-1} = \eta \varepsilon_t.$$

Observe que $x_t; t = 1, 2, 3, \dots, T$ son variables aleatorias independientes que se distribuyen $N(0, \eta^2)$. En términos de la función de densidad, puede escribirse:

$$\begin{aligned} f(x_1; \mu, \rho) &= \frac{1}{(2\pi\eta^2)^{\frac{1}{2}}} \exp\left(\frac{-(y_1 - \mu(1 - \rho) - \rho y_0)^2}{2\eta^2}\right) \\ f(x_2; \mu, \rho) &= \frac{1}{(2\pi\eta^2)^{\frac{1}{2}}} \exp\left(\frac{-(y_2 - \mu(1 - \rho) - \rho y_1)^2}{2\eta^2}\right) \\ &\vdots \\ f(x_T; \mu, \rho) &= \frac{1}{(2\pi\eta^2)^{\frac{1}{2}}} \exp\left(\frac{-(y_T - \mu(1 - \rho) - \rho y_{T-1})^2}{2\eta^2}\right) \end{aligned}$$

Puesto que cada x_t es independiente, puede expresarse la función de densidad conjunta como el producto de sus densidades marginales, es decir,

$$f(x_1, x_2, \dots, x_T) = f(x_1; \mu, \rho) f(x_2; \mu, \rho) \dots f(x_T; \mu, \rho).$$

En consecuencia, la función de verosimilitud está dada por

$$L = f(\rho, \mu | x_t) = \frac{1}{(2\pi\eta^2)^{\frac{T}{2}}} \exp\left(\frac{1}{2\eta^2} - \sum_{t=1}^T (y_t - \mu(1 - \rho) - \rho y_{t-1})^2\right).$$

Si se toma logaritmo natural en ambos lados, se sigue que

$$\log(L) = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\eta^2) - \frac{1}{2\eta^2} \sum_{t=1}^T (y_t - \mu(1 - \rho) - \rho y_{t-1})^2.$$

Ahora considere el problema de optimización:

$$GRAD(\log(L)) = \left\langle \frac{\partial \log(L)}{\partial \mu}, \frac{\partial \log(L)}{\partial \rho} \right\rangle = \vec{0},$$

$$\begin{aligned} \frac{\partial \log(L)}{\partial \mu} &= \frac{1}{\eta^2} 2 \sum_{t=1}^T (y_t - \mu(1 - \rho) - \rho y_{t-1})(1 - \rho) = 0 \\ &\sum_{t=1}^T (y_t - \mu(1 - \rho) - \rho y_{t-1})(1 - \rho) = 0 \\ &\sum_{t=1}^T y_t - \mu(1 - \rho)T - \sum_{t=1}^T \rho y_{t-1} = 0 \\ &\sum_{t=1}^T y_t - \sum_{t=1}^T \rho y_{t-1} = \mu(1 - \rho)T. \end{aligned}$$

Se puede afirmar que para T grande $\frac{1}{T} \sum_{t=1}^T y_t \approx \frac{1}{T} \sum_{t=1}^T y_{t-1}$, por tanto $\frac{1}{T} \sum_{t=1}^T y_t - \rho \frac{1}{T} \sum_{t=1}^T y_t = \mu(1 - \rho) \Rightarrow \frac{1}{T} \sum_{t=1}^T y_t(1 - \rho) = \mu(1 - \rho)$. De este modo $\mu = \frac{1}{T} \sum_{t=1}^T y_t$.

$$\frac{\partial \log(L)}{\partial \rho} = -\frac{1}{\eta^2} 2 \sum_{t=1}^T (y_t - \hat{\mu}(1 - \rho) - \rho y_{t-1}) (\hat{\mu} - y_{t-1}) = 0$$

$$\sum_{t=1}^T (y_t - \hat{\mu}(1 - \rho) - \rho y_{t-1}) (\hat{\mu} - y_{t-1}) = \sum_{t=1}^T -(y_t - \hat{\mu})(y_{t-1} - \hat{\mu}) + \rho(\hat{\mu} - y_{t-1})^2 = 0$$

$$\sum_{t=1}^T \rho(\hat{\mu} - y_{t-1})^2 = \sum_{t=1}^T (y_t - \hat{\mu})(y_{t-1} - \hat{\mu})$$

$$\rho = \frac{\sum_{t=1}^T (y_t - \hat{\mu})(y_{t-1} - \hat{\mu})}{\sum_{t=1}^T (y_{t-1} - \hat{\mu})^2}.$$

Los estimadores de MV para los parámetros μ y ρ son independientes asintóticamente, y equivalentes a

$$\hat{\mu}_T = \frac{1}{T} \sum_{t=1}^T y_t = \bar{y}_T,$$

$$\hat{\rho}_T = \frac{\frac{1}{T} \sum_{t=1}^T (y_t - \bar{y}_T)(y_{t-1} - \bar{y}_T)}{\frac{1}{T} \sum_{t=1}^T (y_{t-1} - \bar{y}_T)^2},$$

donde los residuales son definidos por $\hat{\varepsilon}_t = y_t - \bar{y}_T - \hat{\rho}_T(y_{t-1} - \bar{y}_T)$. Sus varianzas asintóticas están dadas por

$$V_{\text{asy}}(\hat{\mu}_T) = \frac{\eta^2}{T(1 - \rho^2)}, \quad V_{\text{asy}}(\hat{\rho}_T) = \frac{1}{T}(1 - \rho^2).$$

De los estimadores MV de los parámetros μ y ρ , se infieren fácilmente los estimadores MV para los parámetros de interés

$$\hat{\lambda}_T = -\log \hat{\rho}_T, \quad \hat{\phi} = \hat{\lambda} \hat{\mu} \quad \text{y} \quad \hat{\eta} = (\theta(1 - \hat{\rho}^2))^{\frac{1}{2}}.$$

8.2. Simulación

Se realizó un estudio preliminar con una gran cantidad de series simuladas que siguen un proceso con especificación GARCH(1,1) con diferentes valores para los parámetros y con

las restricciones propias para cada uno (ver tabla 8.1). El proceso solución de la Ecuación Diferencial Estocástica (8.1), registra una buena explicación de la dinámica de estas series simuladas cuando $\delta_0 > 0,25$ y el parámetro $\theta = \text{var}[y_t]/E[h_t^2]$.

Parámetros	h_0^2	δ_0	δ_1	θ_1	α	β	y_0
Grafica 1	5	0.5	0.042	0.91	0.9	0.986	12
Grafica 2	7	0.4	0.037	0.819	0.7	0.9	10
Grafica 3	9	0.3	0.034	0.87	0.88	0.97	8
Grafica 4	11	0.29	0.026	0.912	1.2	0.98	6

Figura 8.1: Parámetros de cada simulación

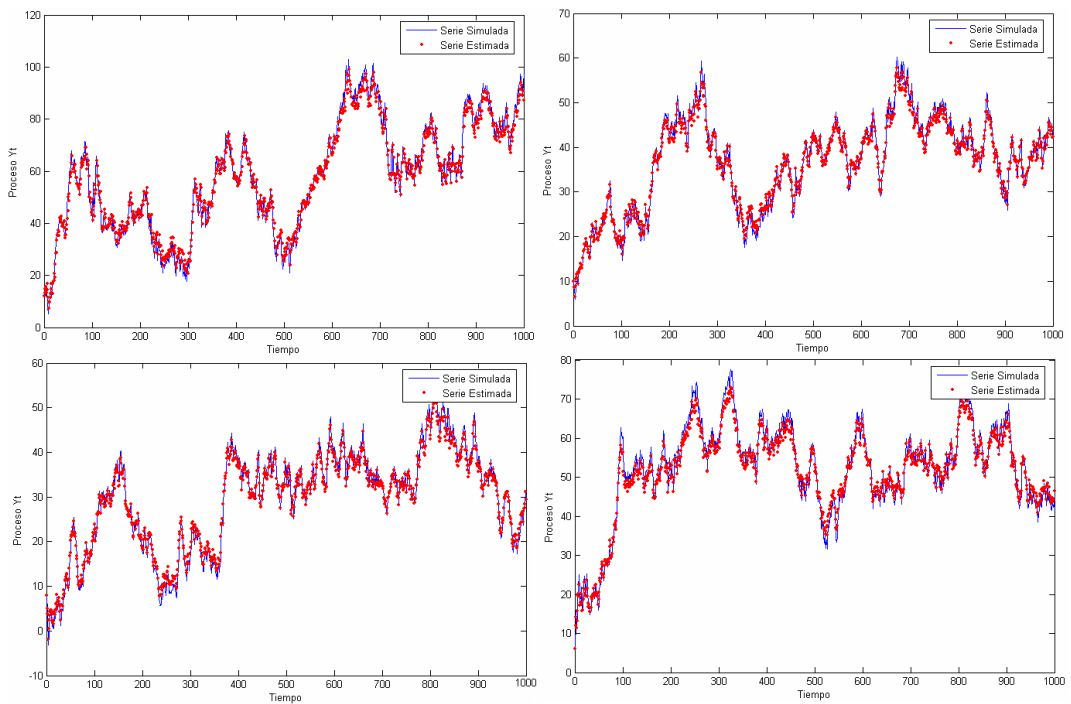


Figura 8.2: Simulaciones para cuatro procesos GARCH

La gráfica 8.2 está organizada de izquierda a derecha y de arriba a abajo; es decir, la gráfica 1 es la superior izquierda y la gráfica 4, la inferior derecha.

8.3. El caso del Aluminio

Como aplicación de esta metodología, se considero el comportamiento dinámico del crecimiento diario de los precios del Aluminio para un total de 1405 datos.

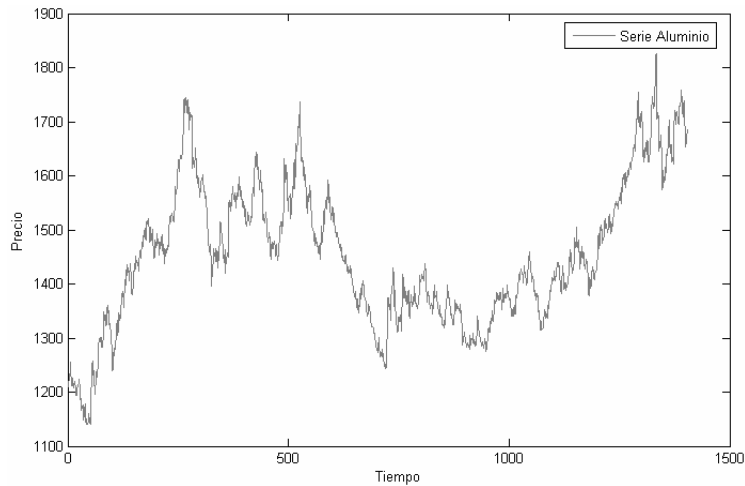


Figura 8.3: Gráfica de la serie de aluminio

Para poder encontrar un modelo apropiado, es necesario que la serie a analizar sea estacionaria, y si no lo es debe diferenciarse las veces necesarias hasta que la serie sea estacionaria en media.

Para encontrar el orden p del modelo $AR(p)$ es necesario observar la autocorrelación parcial y la autocorrelación.

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
██████████	██████████	1 0.991	0.991	1382.3	0.000
██████████	██████████	2 0.982	0.001	2740.3	0.000
██████████	██████████	3 0.974	0.055	4077.4	0.000
██████████	██████████	4 0.966	0.018	5394.6	0.000
██████████	██████████	5 0.958	-0.012	6691.4	0.000
██████████	██████████	6 0.950	-0.008	7967.8	0.000
██████████	██████████	7 0.942	-0.014	9223.5	0.000
██████████	██████████	8 0.935	0.015	10459.0	0.000
██████████	██████████	9 0.926	-0.045	11674.0	0.000
██████████	██████████	10 0.917	-0.065	12865.0	0.000
██████████	██████████	11 0.907	0.014	14032.0	0.000
██████████	██████████	12 0.899	0.008	15178.0	0.000
██████████	██████████	13 0.889	-0.052	16301.0	0.000
██████████	██████████	14 0.879	-0.038	17398.0	0.000
██████████	██████████	15 0.869	0.007	18471.0	0.000
██████████	██████████	16 0.859	0.019	19522.0	0.000

Figura 8.4: Autocorrelación del aluminio

Para determinar el orden del posible modelo $AR(p)$ es necesario mirar los cortes significativos en el PACF. Según los resultados se observa claramente que este proceso es del tipo $AR(1)$ puesto que la correlación presenta un decaimiento exponencial.

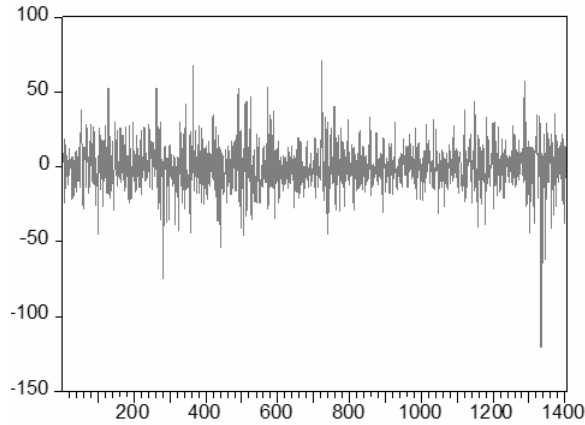


Figura 8.5: Gráfica de los residuales para la serie del aluminio

Para encontrar el mejor modelo, si es el caso, se eliminan los parámetros que son menos significativos, para esto se observa el valor P de cada uno de ellos y se elimina la variable con mayor valor P, y nuevamente se corre el modelo.

Luego de este procedimiento se obtuvo el proceso

$$y_t = 10,44298 + 0,993045y_{t-1} + \varepsilon_t.$$

Dado que este proceso corresponde a un AR(1), se debe observar si este presenta heterocedasticidad condicional. Este análisis resulta de verificar el nuevo correlograma y observar cuales valores de P son menores que 0.05.

Dependent Variable: Y				
Method: ML - ARCH				
Date: 08/15/07 Time: 20:24				
Sample(adjusted): 2 1405				
Included observations: 1404 after adjusting endpoints				
Convergence achieved after 46 iterations				
	Coefficient	Std. Error	z-Statistic	Prob.
C	10.59242	3.583712	2.955712	0.0031
Y(-1)	0.992681	0.002453	404.6471	0.0000
Variance Equation				
C	5.790847	1.632837	3.546494	0.0004
ARCH(1)	0.051336	0.010150	5.057558	0.0000
GARCH(1)	0.925821	0.014680	63.06734	0.0000
R-squared	0.986292	Mean dependent var	1453.113	
Adjusted R-squared	0.986252	S.D. dependent var	132.0643	
S.E. of regression	15.48456	Akaike info criterion	8.252165	
Sum squared resid	335440.3	Schwarz criterion	8.270851	
Log likelihood	-5788.020	F-statistic	25163.81	
Durbin-Watson stat	1.994278	Prob(F-statistic)	0.000000	

Figura 8.6: Modelo GARCH para la serie del aluminio

Después de este análisis se observo un efecto GARCH(1,1) dado por

$$h_t^2 = 5,790847 + 0,051336 \varepsilon_{t-1}^2 + 0,925821 h_{t-1}^2$$

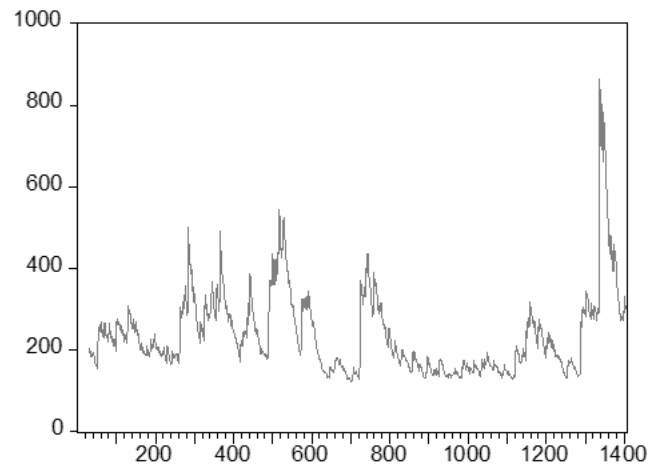
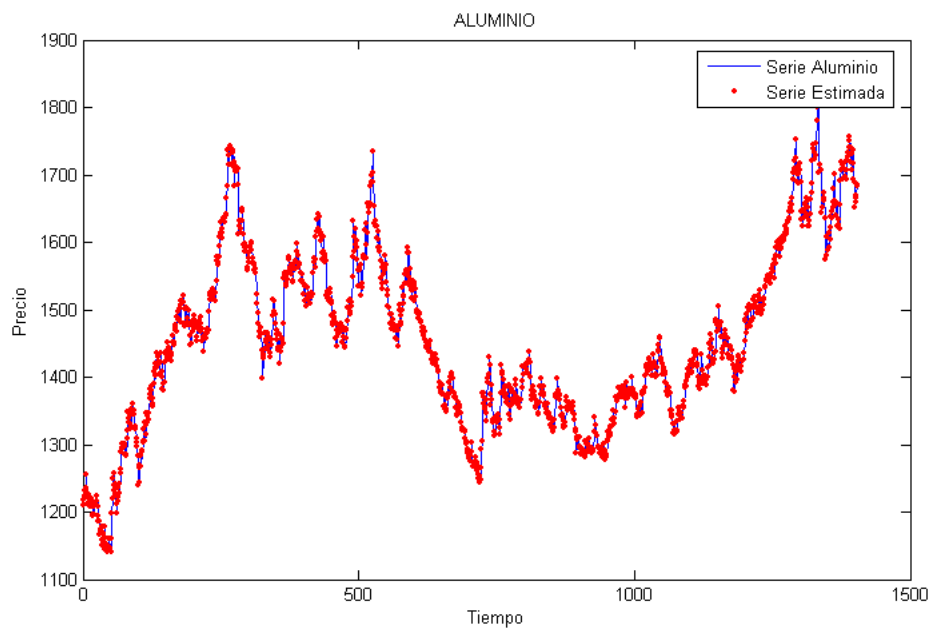


Figura 8.7: Gráfica del modelo GARCH

De este modo se pueden obtener los parámetros de interés δ_0 , δ_1 y θ_1 para ser utilizados en la Ecuación Diferencial Estocástica (8.1), obteniendo así los estimadores de Máxima Verosimilitud $\hat{\rho} = 0,9932$ y $\hat{\eta} = 0,9925$.



8.4. Conclusiones

- Los resultados de este trabajo corroboran una clara relación entre los procesos de reversión a la media del tipo Ornstein Uhlenbeck y los procesos auto-regresivos de orden 1, AR(1).
- Se establece un híbrido entre las series de tiempo no lineales y las Ecuaciones Diferenciales Estocásticas Lineales que dan origen a procesos de reversión a la media con efecto GARCH.
- La implementación del método de Máxima Verosimilitud es muy apropiado para la estimación de parámetros de modelos de difusión, también se han explorado técnicas de algoritmos genéticos y búsqueda tabu que también ofrecen buenos resultados aunque no se tienen claras las propiedades estadísticas de dichos estimadores. Estas técnicas algorítmicas no se incluyeron en este trabajo pero dan pie a investigaciones futuras.
- Este trabajo genera además un panorama de investigación importante en cuanto a la correspondencia entre las series de tiempo, lineales y no lineales, y las Ecuaciones Diferenciales Estocásticas, lo que puede ser una alternativa interesante en el estudio de la dinámica de algunas series de tiempo financieras.

Bibliografía

- [1] Bjork, Tomas (1998). “Arbitrage Theory In Continuous Time”. Oxford University Press Inc. New York.
- [2] Carvajal, Andrés (2003). “Notas de clase sobre equilibrio general”. Universidad del rosario. Departamento de economía, Julio 2003.
- [3] Costa, Luis y Font V, Monserrat (1993). “COMMODITIES Mercados financieros sobre materias primas”. Ed. ESIC, Madrid.
- [4] Cox, J.C., Ingersoll, Jr., J.E., Ross, S. A., “A theory of the term structure of interest rates”, *Econometrica* 53, 385-408., 1985.
- [5] Dixit, A.K. & R.S. Pindyck, “Investment under Uncertainty”, Princeton University Press, Princeton, N.J., 1994.
- [6] Gibson & Schwartz , “Stochastic Convenience Yield and the Pricing of Oil Contingent Claims”, *Journal of Finance*, vol. 45, no 3, pp. 959-976, July 1990.
- [7] Hull, John C. (2003). “Options, futures and other derivatives”. 5a edición. Ed. Prentice Hall. Pearson Educación, New Jersey.
- [8] Lucia, J. J. & Schwartz, E. S., “Electricity prices and power derivatives: Evidence from the Nordic power exchange”, *Review of Derivatives Research* 5(1), 5-50, 2002.
- [9] Mao, Xuerong. (1997). “Stochastic Differential Equations & Aplications”. Horwood Publishing Limited. England.
- [10] Vasicek, O., “An equilibrium characterisation of the term structure”, *Journal of Financial Economics*, 5:-88, 1977.
- [11] Gouriéroux, C. (2001), “Financial econometrics. Problems, models and methods”, Princeton University Press.
- [12] Kloeden, Peter E. and Platen Eckhard (1999). “Numerical Solution of Stochastics Differential Equations”. Springer-Verlag.

- [13] Wilmott, Paul (1998). “Derivatives: The Theory and Practice of Financial Engineering”. John Wiley & Sons.
- [14] Lari-Lavassani, Ali, Sadeghi, Ware, Antony (2001). “Mean Reverting Models For Energy Option Pricing”. Preprint.
- [15] Bachelier, Louis (1900), “Theory of Speculation”, in Paul Cootner, editor: “The Random Character of Stock Market Prices”. (The MIT Press, Cambridge, Mass.)
- [16] Fischer, Black; Myron, Scholes (1973). “The pricing of options and corporates liabilities”. *The journal of political economy*, 81(3): 637-654, may-jun 1973.
- [17] Heston, Steven. “A closed-form solution for options with stochastic volatility with applications to bond a currency options”. *The review of financial studies*, 6(2), 1993.
- [18] Hull, John; White, Alan (1987). “The pricing of options on assets with stochastic volatilities”. *The journal of finance*, 42(2): 281-300, Junio 1987.
- [19] Itô, K. (1944). “Stochastic Integral”. *Proceedings of the Imperial Academy*, Tokyo, 20: 519-524.