



VNIVERSIDAD
D SALAMANCA

Modelos estocásticos en finanzas

Máster Universitario en Modelización Matemática

Facultad de Ciencias

VNIVERSIDAD D SALAMANCA

AUTOR: HÉCTOR ROMERO PAJERO

id00802127@usal.es

TUTOR: ÁNGEL ANDRÉS TOCINO GARCÍA

bacon@usal.es

Salamanca, julio 2025

Trabajo Fin de Máster

Título: Modelos estocásticos en finanzas

Máster Universitario en Modelización Matemática

Facultad de Ciencias

VNiVERSiDAD D SALAMANCA

AUTOR: HÉCTOR ROMERO PAJERO

id00802127@usal.es



TUTOR: ÁNGEL ANDRÉS TOCINO GARCÍA

bacon@usal.es

Salamanca, julio 2025

Resumen

En las últimas décadas, las ecuaciones diferenciales estocásticas (EDEs) se han consolidado como una herramienta clave en las finanzas cuantitativas, al permitir modelar rigurosamente la evolución dinámica e incierta de variables económicas. Este trabajo tiene como objetivo ofrecer una introducción general al uso de EDEs en el ámbito de las finanzas.

Primero, se presentan los fundamentos del cálculo diferencial estocástico, desde el concepto de proceso estocástico, especialmente el movimiento browniano, hasta la fórmula de Itô, pilar central de esta teoría. Luego, se aborda la teoría de las EDEs, incluyendo resultados sobre existencia y unicidad de soluciones, la resolución de ecuaciones lineales y la estabilidad de las soluciones.

Aunque en muchos casos existe solución teórica para una EDE dada, rara vez es posible obtener una expresión analítica explícita. Por este motivo, se introducen los principales métodos numéricos utilizados para resolver EDEs, tales como el esquema de Euler-Maruyama y variantes más avanzadas, los métodos de Taylor estocásticos, predictor-corrector, etc. Se analizan sus propiedades y limitaciones.

Finalmente, se expone la aplicación directa de EDEs en la modelización financiera, explicándose modelos representativos como el modelo de Black-Scholes, el modelo de Heston para volatilidad estocástica, y modelos de tasas de interés como los de Vasicek y Cox-Ingersoll-Ross (CIR), describiendo sus características, así como propiedades de sus soluciones. Además, se proponen métodos numéricos estocásticos diseñados para proporcionar soluciones aproximadas conservando propiedades esenciales de la solución exacta y se presentan simulaciones numéricas ilustrativas.

Abstract

In recent decades, stochastic differential equations (SDEs) have emerged as a cornerstone of quantitative finance, enabling rigorous modeling of the dynamic and uncertain evolution of economic variables. This work provides a comprehensive introduction to the application of SDEs in finance.

First, the foundations of stochastic calculus are presented, ranging from stochastic processes—particularly Brownian motion—to Itô’s formula, the centerpiece of this theory. Subsequently, the theory of SDEs is addressed, including results on the existence and uniqueness of solutions, the analysis of linear equations, and the stability of solutions.

Although theoretical solutions exist for many SDEs, explicit analytical expressions are rarely attainable. Therefore, key numerical methods for solving SDEs are introduced, such as the Euler-Maruyama scheme and other advanced variants (e.g., stochastic Taylor methods, predictor-corrector methods). Their properties and limitations are discussed.

Finally, direct applications of SDEs in financial modeling are demonstrated through representative frameworks, including the Black-Scholes model, Heston’s stochastic volatility model, and interest rate models such as Vasicek and Cox-Ingersoll-Ross (CIR). Their key features and properties of their solutions are described. In addition, stochastic numerical methods—designed to provide approximate solutions while preserving essential properties of the exact solution—are proposed, and illustrative numerical simulations are presented.

Índice

Resumen	i
Abstract	ii
1 Introducción	1
2 Cálculo diferencial estocástico	4
2.1 Preliminares	4
2.2 Procesos de Wiener	6
2.3 La integral de Itô	8
2.4 Fórmula de Itô	14
3 Ecuaciones diferenciales estocásticas	21
3.1 Teorema de existencia y unicidad	21
3.2 Ecuaciones diferenciales estocásticas lineales	25
3.3 Estabilidad de las soluciones	29
4 Métodos numéricos en EDEs	31
4.1 Convergencia de los métodos numéricos	32
4.2 Método de Euler-Maruyama	32
4.3 Método de Milstein	33

4.4	Aproximaciones de Itô-Taylor	34
4.5	Métodos implícitos	38
4.6	Métodos predictor-corrector	41
5	Modelos estocásticos en finanzas	42
5.1	Modelo de Vasicek	42
5.2	Modelo de Black-Scholes	45
5.3	Modelo de salto-difusión	53
5.4	Modelo CIR	55
5.5	Modelos de volatilidad estocástica	62
	Bibliografía	64
	Anexo I: Código Mathematica	66

Índice de figuras

1.0.1 Evolución del precio de cierre del S&P 500.	1
2.2.1 Trayectorias de un proceso de Wiener.	8
3.3.1 Región de estabilidad en media cuadrática en el plano (λ, μ^2)	30
4.5.1 Regiones de estabilidad de los métodos θ de Euler en el plano (λ, μ^2)	40
5.2.1 Opciones europeas a la fecha de vencimiento.	48
5.2.2 Trayectoria del precio del activo subyacente, opción call y opción put.	52
5.3.1 Trayectoria del modelo de salto-difusión de Merton frente a su término difusivo.	54
5.4.1 Trayectoria media y su distancia respecto a la media a largo plazo.	61
5.5.1 Trayectoria del modelo de Heston.	63

Capítulo 1

Introducción

En las últimas décadas, el campo de las finanzas cuantitativas ha experimentado un notable desarrollo gracias a la incorporación de herramientas matemáticas avanzadas. Entre estas herramientas, las *ecuaciones diferenciales estocásticas* (EDEs) han emergido como un pilar fundamental en la modelización de fenómenos financieros complejos. Desde la valoración de derivados financieros hasta la gestión de carteras y el análisis de riesgos, las EDEs permiten describir la evolución dinámica e incierta de variables económicas dentro de un marco riguroso, capturando tanto la tendencia sistemática como las fluctuaciones aleatorias que caracterizan a los mercados financieros.

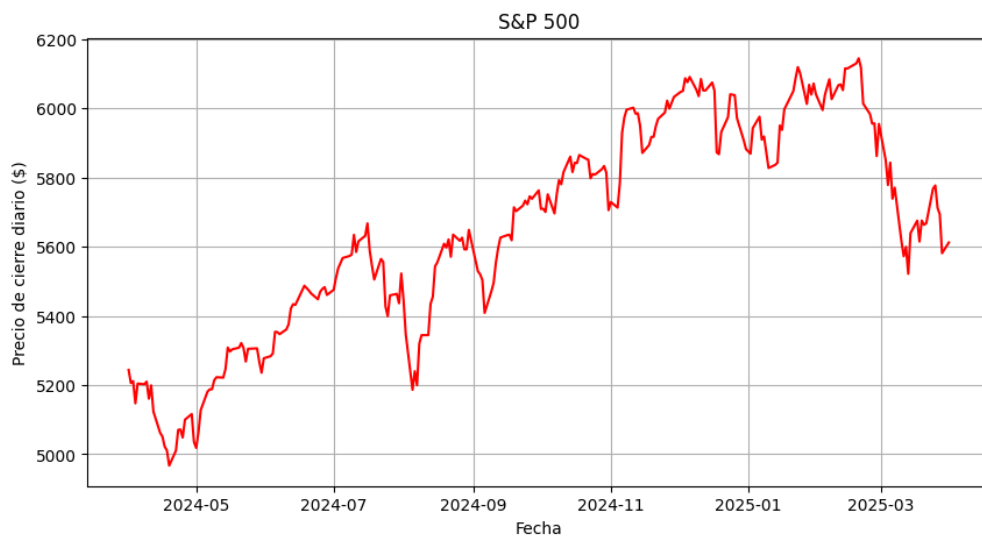


Figura 1.0.1: Evolución del precio de cierre del S&P 500.

Las finanzas, por su naturaleza, están profundamente marcadas por la incertidumbre: los precios de los activos, las tasas de interés, los tipos de cambio y otros indicadores relevantes evolucionan en el tiempo de manera aparentemente errática. La modelización determinista, aunque útil en ciertos contextos, resulta insuficiente para representar fielmente esta aleatoriedad inherente. Es aquí donde las EDEs desempeñan un papel crucial, ya que extienden las ecuaciones diferenciales clásicas al ámbito probabilístico, integrando el ruido aleatorio a través de procesos estocásticos como el movimiento browniano o sus generalizaciones.

El uso de herramientas estocásticas en finanzas tiene sus raíces en el trabajo de Louis Bachelier, quien en su tesis de 1900 introdujo el uso del movimiento browniano para modelar el precio de activos financieros. Aunque sus ideas fueron en gran medida ignoradas por la comunidad matemática de su tiempo, serían redescubiertas y revitalizadas muchas décadas más tarde. El desarrollo riguroso del cálculo estocástico no se consolidaría hasta la obra de Kiyoshi Itô en la década de 1940, quien formuló lo que hoy se conoce como *cálculo de Itô*, estableciendo las bases matemáticas necesarias para trabajar con procesos estocásticos continuos (Itô, 1944).

La consagración del enfoque estocástico en las finanzas llegó con el conocido modelo de Black-Scholes en 1973, que introdujo una EDE para modelar el comportamiento del precio de un activo subyacente y permitió derivar una fórmula cerrada para la valoración de opciones europeas. Este modelo no solo revolucionó los mercados financieros, sino que también dio comienzo a una amplia gama de investigaciones y aplicaciones basadas en EDEs. Desde entonces, se han desarrollado modelos más sofisticados, como los de volatilidad estocástica (por ejemplo, el modelo de Heston), los modelos de tasas de interés como los de Vasicek y Cox-Ingersoll-Ross, y modelos con saltos que incorporan discontinuidades en las trayectorias de precios.

El objetivo de este trabajo es ofrecer una introducción general al uso de ecuaciones diferenciales estocásticas en finanzas. En el primer capítulo se van presentar los fundamentos teóricos del cálculo diferencial estocástico, desde los conceptos básicos de procesos estocásticos hasta la fórmula de Itô. A continuación, se desarrollará una introducción a las ecuaciones diferenciales estocásticas, comentando el teorema de existencia y unicidad de las soluciones, la definición y

propiedades de las EDEs lineales y la estabilidad de las soluciones. En el tercer capítulo, se explican los principales métodos numéricos en EDEs utilizados, mencionando las propiedades propias de cada uno. Finalmente, se presentan algunos de los modelos estocásticos en finanzas más conocidos, en donde se indaga sobre sus características y la solución de las ecuaciones, incluyéndose distintas simulaciones numéricas.

Capítulo 2

Cálculo diferencial estocástico

En esta sección, se va a realizar un repaso de los conceptos fundamentales dentro del cálculo estocástico que permite modelar sistemas y fenómenos que evolucionan de forma no determinística. El principal objeto de trabajo son los procesos estocásticos, que constituyen la base de las ecuaciones diferenciales estocásticas con las que se trabajara más adelante.

2.1 Preliminares

Definición 2.1.1. *Dado $I = [t_0, T]$, un proceso estocástico X sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ es una función medible*

$$\begin{aligned} X : \quad I \times \Omega &\longrightarrow \mathbb{R}^n \\ (t, \omega) &\longmapsto X(t, \omega). \end{aligned}$$

Para cada $t \in [t_0, T]$ fijo,

$$\begin{aligned} X(t) : \quad \Omega &\longrightarrow \mathbb{R}^n \\ \omega &\longmapsto X(t, \omega) \end{aligned}$$

es una variable aleatoria. X_t es una notación equivalente. También se entiende el proceso estocástico como la familia de variables aleatorias $\{X_t\}_{t \in I}$. Análogamente, para cada $\omega \in \Omega$

fijo,

$$\begin{aligned} X(\omega) : T &\longrightarrow \mathbb{R}^n \\ t &\longmapsto X(t, \omega) \end{aligned}$$

es una trayectoria o realización de X . X^ω es una notación equivalente. Esto describe la evolución de un experimento a lo largo del tiempo. Con el paso de tiempo, crece la información conocida de ω .

Definición 2.1.2. Una filtración es una familia $\{\mathcal{F}_t : t \in [t_0, T]\}$ de sub- σ -álgebras de $\mathcal{F}$ tal que $\mathcal{F}_s \subseteq \mathcal{F}_t$ para todo $t_0 \leq s \leq t \leq T$.

Una filtración modela el hecho de que cada vez más información sobre el proceso es conocida con el paso de tiempo.

Definición 2.1.3. Si $\{X_t\}_{t \in I}$ es una familia de variables aleatorias y X_t es $\mathcal{F}_t$ -medible para cada $t \in [t_0, T]$, entonces se dice que $\{X_t\}_{t \in I}$ está adaptado a (o es no-anticipativo con respecto a) la filtración $\{\mathcal{F}_t\}_{t \in I}$.

Para cada $s \in [t_0, t]$, $\sigma\{X_s\}$ es la σ -álgebra generada por X_s , es decir, la σ -álgebra más pequeña sobre Ω que contiene los conjuntos

$$X_s^{-1}(B) \quad \forall B \in \mathcal{B}$$

donde $\mathcal{B}$ denota la σ -álgebra de Borel. Por definición, $\sigma\{X_s\}$ es la σ -álgebra más pequeña donde X_s es medible.

Definición 2.1.4. Un proceso estocástico $\{X_t\}_{t \in I}$ definido sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ es una martingala si es un proceso adaptado con respecto a una filtración $\{\mathcal{F}_t\}_{t \in I}$,

$$\mathbb{E}[|X_t|] < \infty \quad \forall t \in I$$

y

$$\mathbb{E}[X_t | \mathcal{F}_s] = X_s \quad \text{casi seguro (c.s.),} \quad \forall s \leq t \in I.$$

Una martingala, denotada normalmente por la pareja $\{X_t, \mathcal{F}_t\}_{t \in I}$, es continua si existe un conjunto $\Omega_0 \subset \Omega$ con probabilidad 1 (c.p.t) tal que la trayectoria $t \mapsto X_t(\omega)$ es continua $\forall \omega \in \Omega_0$.

Dado un proceso estocástico $\{X_t\}_{t \in I}$ unidimensional o escalar

$$X : (I \times \Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

para cada conjunto finito $\{t_1, \dots, t_n\} \subset I$ la variable vectorial

$$(X_{t_1}, \dots, X_{t_n}) : \Omega \rightarrow \mathbb{R}^n$$

induce la distribución $P_{t_1, \dots, t_n}$ definida por

$$P_{t_1, \dots, t_n}(B) = P(\omega \in \Omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B)$$

que determina una función de distribución finito-dimensional

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n)$$

Al variar $n \in \mathbb{N}, t_1, \dots, t_n \in [t_0, T]$ la familia de distribuciones satisface las propiedades de simetría y compatibilidad. Recíprocamente, dada una familia $\{F_{t_1, \dots, t_n}\}_{n \in \mathbb{N}, t_1, \dots, t_n \in [t_0, T]}$ que cumple estas propiedades, existe un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y un proceso aleatorio $\{X_t\}_{t \in [t_0, T]}$ con distribuciones finito-dimensionales $\{F_{t_1, \dots, t_n}\}$.

2.2 Procesos de Wiener

Definición 2.2.1. *Un proceso estocástico $\{X_t\}_{t \in I}$ es un proceso gaussiano si sus distribuciones finito-dimensionales $\{F_{t_1, \dots, t_n}\}$ son distribuciones normales.*

Un proceso Gaussiano $\{X_t\}$ está definido completamente por su media,

$$\begin{aligned} m_X : I &\longrightarrow \mathbb{R} \\ t &\longmapsto \mathbb{E}[X_t] \end{aligned}$$

y su covarianza,

$$\begin{aligned} \text{cov}(X_{t_1}, X_{t_2}) &= \mathbb{E}[(X_{t_1} - \mathbb{E}[X_{t_1}])(X_{t_2} - \mathbb{E}[X_{t_2}])] \\ &= \mathbb{E}[X_{t_1}X_{t_2}] - \mathbb{E}[X_{t_1}]\mathbb{E}[X_{t_2}] \end{aligned}$$

Definición 2.2.2. *Un proceso de recuento es un proceso estocástico $\{N_t\}_{t \in I}$ de valor entero, cuyas variables aleatorias contabilizan el número de sucesos aleatorios ocurridos en un intervalo de tiempo continuo.*

Definición 2.2.3. *Un proceso de recuento $\{N_t\}_{t \geq 0}$ se denomina proceso de Poisson con parámetro λ si satisface:*

1. $N_0 = 0$.
2. Posee incrementos independientes, es decir,

$$N_{t_1} - N_{t_0}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$$

son independientes para todo $0 \leq t_0 < t_1 < \dots < t_n$.

3. $N_t - N_s \sim \text{Poisson}(\lambda(t - s))$, para cualquier $t > s \geq 0$, con $\lambda > 0$.

A partir de esta idea, se tiene que los incrementos son estacionarios, ya que la distribución de los incrementos depende únicamente de la longitud del intervalo y no de su posición temporal. Además, el número medio de sucesos ocurridos en un intervalo de longitud t es λt . Por ello, el parámetro λ se llama razón (o intensidad) del proceso, el número de sucesos por unidad de tiempo.

Definición 2.2.4. *Un proceso estocástico $\{W_t\}_{t \geq 0}$ es un proceso de Wiener estándar o movimiento browniano si cumple las siguientes propiedades:*

1. $W_0 = 0$ c.p.t.
2. $\{W_t\}$ es un proceso gaussiano, con $W_t \sim \mathcal{N}(0, t)$ y $\text{cov}(W_t, W_s) = \min\{t, s\}$
3. $\{W_t\}$ tiene incrementos independientes, o equivalentemente,

$$W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$$

son independientes para todo $n > 1$ y $0 \leq t_0 < t_1 < \dots < t_n$.

Se demuestra que para cualquier proceso de Wiener $\{W_t\}$ se cumplen las propiedades

- $W_t - W_s \sim \mathcal{N}(0, t - s)$ para cualquier $0 \leq s < t$.
- Existe $\tilde{\Omega} \subset \Omega$, c.p.t, tal que $t \mapsto W_t(\omega)$ es continuo para casi todo $\omega \in \tilde{\Omega}$.

Un proceso de Wiener d -dimensional es un proceso estocástico vectorial

$$W_t = (W_t^1, W_t^2, \dots, W_t^d)$$

cuyos componentes W_t^i son procesos de Wiener unidimensionales independientes, y

$$W_t - W_s \sim \mathcal{N}(0, (t - s)I_d)$$

donde I_d es la matriz identidad de dimensiones $d \times d$.

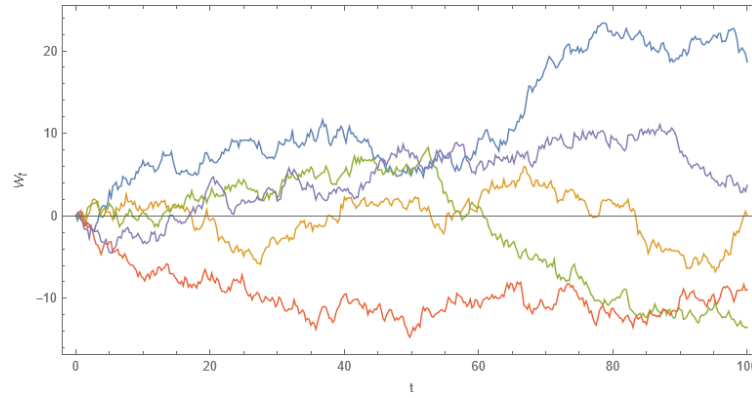


Figura 2.2.1: Trayectorias de un proceso de Wiener.

2.3 La integral de Itô

En esta sección, se trata la construcción de la integral de Itô. Primero, se deben introducir el conjunto de funciones para las cuales esta integral tiene sentido.

Definición 2.3.1. Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, un proceso estocástico $f : [t_0, T] \times \Omega \rightarrow \mathbb{R}^n$ se dice que es cuadrado integrable si

$$\|f\|_{(2, [t_0, T])} = \left(\int_{t_0}^T \mathbb{E}[|f_t|^2] dt \right)^{1/2} < +\infty$$

Esto se denota con $f \in \mathcal{L}^2([t_0, T] \times \Omega; \mathbb{R}^n)$. A su vez, f es integrable y se escribe $f \in \mathcal{L}^1([t_0, T] \times \Omega; \mathbb{R}^n)$ si

$$\|f\|_{(1, [t_0, T])} = \left(\int_{t_0}^T \mathbb{E}[|f_t|] dt \right) < +\infty$$

Entonces, la integral de Itô se definirá en el conjunto denominado $\mathcal{M}^2[t_0, T]$ que engloba la clase de funciones $f = f(t, \omega)$

$$f : [t_0, T] \times \Omega \rightarrow \mathbb{R}$$

con las siguientes propiedades:

1. f es una función $(\mathcal{B}[t_0, T] \times \mathcal{F})$ -medible.
2. $f \in \mathcal{L}^2([t_0, T] \times \Omega; \mathbb{R})$.
3. f está adaptado a una filtración $\{\mathcal{F}_t\}$, i.e. $f(t, \cdot)$ es $\mathcal{F}_t$ -medible para cada $t \in [t_0, T]$.

Primero, se va a definir la integral estocástica para funciones escalonadas.

Definición 2.3.2. Una función $g \in \mathcal{M}^2[t_0, T]$ es una función escalonada si existe una descomposición $t_0 < t_1 < t_2 < \dots < t_n = T$ de forma que $g(s) = g(t_{i-1})$, $\forall s \in [t_{i-1}, t_i]$, es decir,

$$g(s) = \sum_{i=1}^n g(t_{i-1}) \chi_{[t_{i-1}, t_i)}(s)$$

donde

$$\chi_{[c, d]}(s) = \begin{cases} 1 & \text{si } s \in [c, d] \\ 0 & \text{en cualquier otro caso} \end{cases}$$

es la función característica del intervalo $[c, d]$.

Para una función escalonada g se define

$$I[g] = \int_{t_0}^T g(s) dW_s = \sum_{i=1}^n g(t_{i-1})(W_i - W_{i-1}).$$

El conjunto de estas funciones escalonadas se denota $\mathcal{S}^2[t_0, T] \subseteq \mathcal{M}^2[t_0, T]$. Entonces, el objetivo se trata de aproximar cualquier $f \in \mathcal{M}^2[t_0, T]$ por una sucesión de funciones escalonadas $\{f_n\}_{n \in \mathbb{N}}$ y definir la integral de Itô de f como el límite de las integrales de dichas funciones escalonadas, que ya son conocidas. Para ello, se debe demostrar el siguiente lema.

Lema 2.3.3. $S^2[t_0, T]$ es denso en $(\mathcal{M}^2[t_0, T], \|\cdot\|_{(2,T)})$.

Demostración. Considérense particiones del intervalo $[t_0, T]$ de la forma $t_0^{(n)} < t_1^{(n)} < \dots < t_n^{(n)}$ con $t_{j+1}^{(n)} - t_j^{(n)} \rightarrow 0$ para $j = 0, 1, \dots, n$, con $n \rightarrow +\infty$. Para f continua en media cuadrática, se define una sucesión de funciones escalonadas $\{f_n\}_{n \in \mathbb{N}} \subset S_T$ con $f_n(t, \omega) = f(t_j^{(n)}, \omega)$, con probabilidad 1, $t \in [t_j^{(n)}, t_{j+1}^{(n)})$ para $j = 0, 1, \dots, n$. Claramente,

$$\mathbb{E}[|f_n(t, \cdot) - f(t, \cdot)|^2] \rightarrow 0 \text{ cuando } n \rightarrow +\infty$$

para cada $t \in [0, T]$. Luego, por el teorema de convergencia dominada de Lebesgue, se tiene que

$$\lim_{n \rightarrow +\infty} \int_{t_0}^T \mathbb{E}[|f_n(t, \cdot) - f(t, \cdot)|^2] dt = 0$$

es decir, $\|f_n - f\|_{(2,T)} \rightarrow 0$ cuando $n \rightarrow +\infty$, siendo este el resultado al que se quiere llegar. La cuestión es que una función $f \in \mathcal{M}^2[t_0, T]$ no es generalmente continua en media cuadrática, pero se puede aproximar para la norma $\|\cdot\|_{(2,T)}$ por una que sea. Primero, se debe aproximar f con una función acotada $f_N \in \mathcal{M}^2[t_0, T]$ definida por:

$$f_N(t, \omega) = \max\{-N, \min\{f(t, \omega), N\}\}$$

para algún $N > 0$. $|f_N(t, \omega)| \leq N$, con $f_N(t, \omega) = f(t, \omega)$ cuando $|f(t, \omega)| \leq N$. Además,

$$\int_{t_0}^T \mathbb{E}[|f_N(t, \cdot) - f(t, \cdot)|^2] dt \leq 4 \int_{t_0}^T \mathbb{E}[|f(t, \cdot)|^2] dt < +\infty$$

luego por el teorema de convergencia dominada aplicado a las funciones $\mathbb{E}[|f_N(t, \cdot) - f(t, \cdot)|^2]$ se concluye que

$$\lim_{N \rightarrow +\infty} \int_{t_0}^T \mathbb{E}[|f_N(t, \cdot) - f(t, \cdot)|^2] dt = 0.$$

Ahora, para dicha f_N , se define una función suavizada g_k , para $k > 0$, mediante

$$g_k(t, \omega) = ke^{-kt} \int_{t_0}^t e^{ks} f_N(s, \omega) ds.$$

Por las propiedades de f_N y el hecho de que la anterior integral no involucra valores de f_N con valores que excedan t , se tiene que g_k es una función $(\mathcal{B} \times \mathcal{F})$ -medible y $g_k(t, \cdot)$ es $\mathcal{F}_t$ -medible para cada $t \in [0, T]$. Además, dado que $|f_N| \leq N$, se consigue

$$|g_k(t, \omega)| \leq ke^{-kt} \int_{t_0}^t e^{ks} N ds = Nke^{-kt} \frac{e^{kt} - 1}{k} = N(1 - e^{-kt}) \leq N$$

de modo que $\mathbb{E}[g_k^2]$ es finito y acotado por N^2 , e integrable sobre $0 \leq t \leq T$, luego $g_k \in \mathcal{M}^2[t_0, T]$. Finalmente, se satisface que

$$|g_k(t, \omega) - g_k(s, \omega)| \leq 2kN|t - s|$$

y, por tanto, g_k es continua en t para cada ω fijo. Esta cota implica que $\mathbb{E}[g_k^2]$ es continua, es decir, g_k es continua en media cuadrática. Por consiguiente, se puede aproximar por una sucesión de funciones escalonadas $\{f_n\}_{n \in \mathbb{N}} \in S_T$ como se hizo anteriormente. Para cualquier $\epsilon > 0$ se puede escoger una f_N , g_k y $\{f_n\}_{n \in \mathbb{N}}$ de forma que

$$\|f - f_N\|_{(2,T)} \leq \frac{1}{3}\epsilon, \quad \|f_N - g_k\|_{(2,T)} \leq \frac{1}{3}\epsilon, \quad \|g_k - f_n\| \leq \frac{1}{3}\epsilon.$$

Luego, por la desigualdad triangular, se tiene

$$\|f - f^{(n)}\|_{(2,T)} \leq \|f - f_N\|_{(2,T)} + \|f_N - g_k\|_{(2,T)} + \|g_k - f_n\|_{(2,T)} < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Por lo tanto, f_n es una función escalonada que aproxima f dentro de ϵ , lo que demuestra que S_T es denso en $\mathcal{M}^2[t_0, T]$. $\square$

Lema 2.3.4. Para cualquier $f, g \in S^2[t_0, T]$ y $\alpha, \beta \in \mathbb{R}$ la integral de Itô definida por

$$I[f](\omega) = \sum_{i=1}^n f_{t_{i-1}}(W_i - W_{i-1})$$

satisface:

1. $I[f]$ es $\mathcal{F}_T$ -medible.
2. $\mathbb{E}[I[f]] = 0$.
3. $\mathbb{E}[I[f]^2] = \int_{t_0}^T \mathbb{E}[f_t^2] dt$ (Isometría de Itô).
4. $I[\alpha f + \beta g] = \alpha I[f] + \beta I[g]$, c.p.t.

Demostración.

1. Como $f_{t_{j-1}}$ es $\mathcal{F}_{t_{j-1}}$ -medible y $(W_{t_j} - W_{t_{j-1}})$ es $\mathcal{F}_{t_j}$ -medible, donde $\mathcal{F}_{t_j} \subseteq \mathcal{F}_{t_{j-1}}$, su producto $f_{t_j}(W_{t_j} - W_{t_{j-1}})$ es $\mathcal{F}_{t_j}$ -medible para $j = 0, 1, 2, \dots, n$. Por tanto, $I[f]$ es $\mathcal{F}_T$ -medible.

2. Dado que f es adaptado a $\{\mathcal{F}_t\}$, $f_{t_{j-1}}$ es $\mathcal{F}_{t_{j-1}}$ -medible y $(W_{t_j} - W_{t_{j-1}})$ es independiente de $\mathcal{F}_{t_{j-1}}$ y se conoce que $(W_t - W_s) \sim \mathcal{N}(0, t - s)$ para cualquier $0 \leq s < t$. Luego,

$$\mathbb{E}[I[f]] = \sum_{i=1}^n \mathbb{E} \left[f_{t_{i-1}} (W_i - W_{i-1}) \right] = \sum_{i=1}^n \mathbb{E}[f_{t_{i-1}}] \mathbb{E}[(W_i - W_{i-1})] = 0.$$

3. Sabiendo que $(W_t - W_s) \sim \mathcal{N}(0, t - s)$ y para cualquier $0 \leq s < t$,

$$\begin{aligned} \mathbb{E}[I[f]^2] &= \sum_{j=1}^n \mathbb{E} \left[f_{t_{j-1}}^2 (W_{t_j} - W_{t_{j-1}})^2 \right] \\ &\quad + 2 \sum_{j=1}^n \sum_{i=j+1}^n \mathbb{E} \left[f_{t_j} f_{t_i} (W_{t_j} - W_{t_{j-1}})(W_{t_i} - W_{t_{i-1}}) \right] \\ &= \sum_{j=1}^n \mathbb{E}[f_{t_j}^2] (t_j - t_{j-1}) = \int_0^T \mathbb{E}[f_t^2] dt. \end{aligned}$$

4. Sea $\alpha f + \beta g \in S^2[t_0, T]$, con la combinación de los puntos escalonados de f y g , se tiene que

$$\begin{aligned} I[\alpha f + \beta g] &= \sum_{i=1}^n (\alpha f + \beta g)_{t_i} (W_{t_i} - W_{t_{i-1}}) \\ &= \alpha \sum_{i=1}^n f_{t_i} (W_{t_i} - W_{t_{i-1}}) + \beta \sum_{i=1}^n g_{t_i} (W_{t_i} - W_{t_{i-1}}) \\ &= \alpha I[f] + \beta I[g] \end{aligned}$$

Lema 2.3.5. Sea $f \in \mathcal{M}^2[t_0, T]$ y $\{f_n\}_{n \in \mathbb{N}}$ una sucesión de funciones escalonadas que converge a f . Entonces existe una única variable aleatoria $I \in \mathcal{L}^2(\Omega; \mathbb{R})$, c.p.t, tal que

$$\lim_{n \rightarrow \infty} \mathbb{E}[|I[f_n] - I|^2] = 0.$$

Además I no depende de la sucesión de funciones escalonadas que converge a f .

Demostración. Para cada $n \in \mathbb{N}$ la integral de Itô $I[f_n]$ está definida por lema 2.3.4, y, como $I[f_n] - I[f_{n+m}] = I[f_n - f_{n+m}]$, se satisface que

$$\begin{aligned} \mathbb{E} \left[\left| \int_0^T f_n(t, \omega) dW_t - \int_0^T f_{n+m}(t, \omega) dW_t \right|^2 \right] &= \mathbb{E} \left[\int_0^T |f_n(t, \omega) - f_{n+m}(t, \omega)|^2 dt \right] \\ &\leq 2 \int_0^T \left(\mathbb{E} [|f_n(t, \omega) - f(t, \omega)|^2] + \mathbb{E} [|f_{n+m}(t, \omega) - f(t, \omega)|^2] \right) dt, \end{aligned}$$

donde se utilizó que $(a + b)^2 \leq 2(a^2 + b^2)$. Esta expresión converge a 0 cuando $n \rightarrow +\infty$,

luego

$$\left\{ \int_0^T f_n(t, \omega) dW_t \right\}_{n \in \mathbb{N}}$$

es una sucesión de Cauchy en el espacio $\mathcal{L}^2(\Omega; \mathbb{R})$. Por tanto, existe y es única, c.p.t, la variable aleatoria I tal que $\mathbb{E}[|I[f_n] - I|^2] \rightarrow 0$ cuando $n \rightarrow +\infty$. Además, I es $\mathcal{F}_T$ -medible ya que es el límite de variables aleatorias $\mathcal{F}_T$ -medibles.

Ahora, se considera $\{\bar{f}_n\}$ otra sucesión de funciones escalonadas que converge a f y supóngase que $I[\bar{f}_n]$ converge a $\bar{I}$. Entonces,

$$\mathbb{E}[|I - \bar{I}|^2] \leq 2 \left(\mathbb{E}[|I - I[f_n]|^2] + \mathbb{E}[|\bar{I} - I[\bar{f}_n]|^2] \right)$$

donde se ha estimado de la misma forma de antes. Tomando límites, se obtiene que $\mathbb{E}[|I - \bar{I}|^2] = 0$ cuando $n \rightarrow +\infty$, luego se demuestra $I = \bar{I}$, c.p.t. $\square$

Con todo esto, se tiene lo necesario para definir la integral de Itô.

Definición 2.3.6. Dada una función $f \in \mathcal{M}^2[t_0, T]$, se define su integral de Itô $I[f]$ como el límite común en media cuadrática de las secuencias de las sumas de cualquier sucesiones de funciones escalonadas $f_n \in S^2[t_0, T]$ convergentes a f por la norma $\|\cdot\|_{(2,T)}$, i. e.

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\left(I[f] - \int_0^T f_n(s) dW_s \right)^2 \right] = 0.$$

Claramente, esta definición hereda las propiedades nombradas en el lema 2.3.4.

Teorema 2.3.7. La integral de Itô $I[f]$ para funciones $f \in \mathcal{M}^2[t_0, T]$ satisface:

1. $I[f]$ es $\mathcal{F}_T$ -medible.
2. $\mathbb{E}[I[f]] = 0$.
3. $\mathbb{E}[I[f]^2] = \int_{t_0}^T \mathbb{E}[f_t^2] dt$ (Isometría de Itô).
4. $I[\alpha f + \beta g] = \alpha I[f] + \beta I[g]$, c.p.t.

Para un subintervalo $[t_1, t] \subseteq [t_0, T]$, se construye el proceso $Z = \{Z_t, t_1 \leq t \leq T\}$ definido por

$$Z_t = \int_{t_1}^t f_s dW_s$$

c.p.t, para $t_1 \leq t \leq T$. Se verifica fácilmente que Z_t es $\mathcal{F}_t$ -medible con $\mathbb{E}[Z_t] = 0$ y

$$\mathbb{E}[Z_t^2] = \int_{t_1}^t \mathbb{E}[f_s^2] ds.$$

Dado un proceso de Wiener d -dimensional $W_t = (W_t^1, W_t^2, \dots, W_t^d)$ y

$$f = \begin{pmatrix} f_{11} & \cdots & f_{1d} \\ \vdots & \ddots & \vdots \\ f_{n1} & \cdots & f_{nd} \end{pmatrix} : [t_0, T] \times \Omega \rightarrow \mathbb{R}^{n \times d}$$

se define

$$\int_{t_0}^T f(t, \omega) dW_t = \left(\sum_{i=1}^d \int_{t_0}^T f_{1i}(t, \omega) dW_t^i, \dots, \sum_{i=1}^d \int_{t_0}^T f_{ni}(t, \omega) dW_t^i \right).$$

2.4 Fórmula de Itô

Definición 2.4.1. Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, un proceso estocástico $\{X_t\}_{t \in T}$ se denomina proceso de Itô si es de la forma

$$X_t = c + \int_{t_0}^t f(s) ds + \int_{t_0}^t g(s) dW_s,$$

c.p.t, donde $f(t, \omega) \in \mathcal{L}^1[t_0, T]$ se denomina coeficiente de deriva, $g(t, \omega) \in \mathcal{M}^2[t_0, T]$ se denomina coeficiente de difusión y $c(\omega)$ una variable aleatoria $\mathcal{F}_{t_0}$ -medible.

Otra forma de expresarlo es

$$dX_t = f(t)dt + g(t)dW_t$$

Teorema 2.4.2 (Fórmula de Itô unidimensional). Sea

$$V : [t_0, T] \times \mathbb{R} \longrightarrow \mathbb{R}$$

una función con derivadas parciales $\frac{\partial V}{\partial t}$, $\frac{\partial V}{\partial x}$, $\frac{\partial^2 V}{\partial x^2}$ continuas, y $\{X_t\}$ un proceso de Itô unidimensional que verifica la relación

$$dX_t = f(t)dt + g(t)dW_t.$$

Entonces, el proceso $Y_t = V(t, X_t)$ verifica

$$dY_t = \left(\frac{\partial V}{\partial t}(t, X_t) + \frac{\partial V}{\partial x}(t, X_t)f(t) + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(t, X_t)g^2(t) \right) dt + \frac{\partial V}{\partial x}(t, X_t)g(t)dW_t,$$

que también puede expresarse

$$dY_t = \frac{\partial V}{\partial t}(t, X_t) dt + \frac{\partial V}{\partial x}dX_t + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(t, X_t)g^2(t).$$

Demostración. Se asume que $f(t)$ y $g(t)$ son funciones escalonadas y que las derivadas parciales $\frac{\partial V}{\partial t}$, $\frac{\partial V}{\partial x}$, $\frac{\partial^2 V}{\partial t^2}$, $\frac{\partial^2 V}{\partial x^2}$, $\frac{\partial^2 V}{\partial t \partial x}$ son continuas y acotadas. Se toma una partición del intervalo $[t_0, t]$ con $t_0 < t_1 < \dots < t_N = t$ y se define:

$$\Delta t_i = t_{i+1} - t_i, \quad \Delta X_i = X_{t_{i+1}} - X_{t_i}, \quad \Delta W_i = W_{t_{i+1}} - W_{t_i}$$

Como X_t es un proceso de Itô, en cada subintervalo $[t_i, t_{i+1}]$:

$$\Delta X_i = \int_{t_i}^{t_{i+1}} f(s) ds + \int_{t_i}^{t_{i+1}} g(s) dW_s.$$

Dado que f y g son funciones escalonadas, se tiene entonces:

$$\Delta X_i = f(t_i)\Delta t_i + g(t_i)\Delta W_i.$$

Como $dX_t = f(t)dt + g(t)dW_t$, en cada partición:

$$\Delta X_i = f(t_i)\Delta t_i + g(t_i)\Delta W_i + \text{términos pequeños}$$

Con ello, se aplica el desarrollo de Taylor de segundo orden a V en (t_i, X_{t_i}) para $V(t_{i+1}, X_{t_{i+1}})$:

$$\begin{aligned} V(t_{i+1}, X_{t_{i+1}}) = & V(t_i, X_{t_i}) + \frac{\partial V}{\partial t}\Delta t_i + \frac{\partial V}{\partial x}\Delta X_i + \frac{\partial^2 V}{\partial t^2}(\Delta t_i)^2 \\ & + \frac{\partial^2 V}{\partial t \partial x}\Delta t_i\Delta X_i + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(\Delta X_i)^2 + R_i \end{aligned}$$

donde R_i es un resto que incluye términos de orden superior y las derivadas se evalúan en (t_i, X_{t_i}) . Considerando $Y_t = V(t, X_t)$, se construye

$$Y_t - Y_{t_0} = \sum_{i=0}^{N-1} [V(t_{i+1}, X_{t_{i+1}}) - V(t_i, X_{t_i})].$$

Sustituyendo en la expansión de Taylor:

$$\begin{aligned} Y_t - Y_{t_0} = & \sum_i \frac{\partial V}{\partial t} \Delta t_i + \sum_i \frac{\partial V}{\partial x} \Delta X_i + \sum_i \frac{\partial^2 V}{\partial t^2} (\Delta t_i)^2 \\ & + \sum_i \frac{\partial^2 V}{\partial t \partial x} \Delta t_i \Delta X_i + \frac{1}{2} \sum_i \frac{\partial^2 V}{\partial x^2} (\Delta X_i)^2 + \sum_i R_i \end{aligned}$$

Teniendo este resultado, se necesita analizar el comportamiento de cada término cuando $N \rightarrow +\infty$, es decir, cuando la partición tiende a cero. Del primer término se tiene

$$\sum_i \frac{\partial V}{\partial t} \Delta t_i \xrightarrow{N \rightarrow +\infty} \int_{t_0}^t \frac{\partial V}{\partial t}(s, X_s) ds$$

que es la convergencia a la integral de Riemann. Del segundo término, sustituyendo $\Delta X_i = f(t_i)\Delta t_i + g(t_i)\Delta W_i$,

$$\sum_i \frac{\partial V}{\partial x} \Delta X_i = \sum_i \frac{\partial V}{\partial x} (f(t_i)\Delta t_i + g(t_i)\Delta W_i).$$

Al tomar límite

$$\begin{aligned} \sum_i \frac{\partial V}{\partial x} f(t_i) \Delta t_i & \rightarrow \int_{t_0}^t \frac{\partial V}{\partial x} f(s) ds \\ \sum_i \frac{\partial V}{\partial x} g(t_i) \Delta W_i & \rightarrow \int_{t_0}^t \frac{\partial V}{\partial x} g(s) dW_s \end{aligned}$$

donde la primera parte es la convergencia a la integral de Riemann y la segunda a la integral de Itô. Del tercer término,

$$\sum_i \frac{\partial^2 V}{\partial t^2} (\Delta t_i)^2 \rightarrow 0$$

debido a la acotación y continuidad de $\frac{\partial^2 V}{\partial t^2}$. En cuanto al cuarto término, sustituyendo ΔX_i , se desarrolla:

$$\sum_i \frac{\partial^2 V}{\partial t \partial x} \Delta t_i \Delta X_i = \underbrace{\sum_i \frac{\partial^2 V}{\partial t \partial x} f(t_i) (\Delta t_i)^2}_{(I)} + \underbrace{\sum_i \frac{\partial^2 V}{\partial t \partial x} g(t_i) \Delta t_i \Delta W_i}_{(II)}.$$

Para (I), como $\frac{\partial^2 V}{\partial t \partial x}$ y f son acotados, se tiene:

$$\left| \sum_i \frac{\partial^2 V}{\partial t \partial x} f(t_i) (\Delta t_i)^2 \right| \leq C \sum_i (\Delta t_i)^2 \leq C \max_j (\Delta t_j) \sum_i \Delta t_i.$$

Cuando $N \rightarrow +\infty$, entonces $\max_j(\Delta t_j) \rightarrow 0$. Luego,

$$\sum_i \frac{\partial^2 V}{\partial t \partial x} f(t_i) (\Delta t_i)^2 \xrightarrow{\mathcal{L}^2} 0.$$

Para (II), se desarrolla con la norma $\mathcal{L}^2$:

$$\begin{aligned} \mathbb{E} \left[\left(\sum_i \frac{\partial^2 V}{\partial t \partial x} g(t_i) \Delta t_i \Delta W_i \right)^2 \right] &= \sum_i \mathbb{E} \left[\left(\frac{\partial^2 V}{\partial t \partial x} g(t_i) \Delta t_i \right)^2 \right] \mathbb{E}[(\Delta W_i)^2] \\ &= \sum_i \mathbb{E} \left[\left(\frac{\partial^2 V}{\partial t \partial x} g(t_i) \Delta t_i \right)^2 \right] \Delta t_i. \end{aligned}$$

Por continuidad y acotación,

$$\mathbb{E} \left[\left(\frac{\partial^2 V}{\partial t \partial x} f(t_i) g(t_i) \Delta t_i \right)^2 \right] \leq K^2 (\Delta t_i)^2.$$

Así que:

$$\sum_i \mathbb{E} \left[\left(\frac{\partial^2 V}{\partial t \partial x} f(t_i) g(t_i) \Delta t_i \right)^2 \right] \Delta t_i \leq K^2 \sum_i (\Delta t_i)^3 \leq K^2 \max_j(\Delta t_j) \sum_i (\Delta t_i)^2 \rightarrow 0.$$

Por tanto,

$$\sum_i \frac{\partial^2 V}{\partial t \partial x} f(t_i) g(t_i) \Delta t_i \Delta W_i \xrightarrow{\mathcal{L}^2} 0.$$

Para el quinto término, sustituyendo ΔX_i , se desarrolla

$$\frac{1}{2} \sum_i \frac{\partial^2 V}{\partial x^2} (\Delta X_i)^2 = \frac{1}{2} \sum_i \frac{\partial^2 V}{\partial x^2} \underbrace{[f^2(t_i) (\Delta t_i)^2]}_{(A)} + \underbrace{2f(t_i) g(t_i) \Delta t_i \Delta W_i}_{(B)} + \underbrace{g^2(t_i) (\Delta W_i)^2}_{(C)}$$

Analizando cada componente:

- Con (A), análogo a (I), se tiene:

$$\left| \frac{\partial^2 V}{\partial x^2} f^2(t_i) (\Delta t_i)^2 \right| \leq C \sum_i (\Delta t_i)^2 \leq C \max_j(\Delta t_j) \sum_i \Delta t_i.$$

Cuando $N \rightarrow +\infty$, entonces $\max_j(\Delta t_j) \rightarrow 0$. Luego,

$$\frac{\partial^2 V}{\partial x^2} f^2(t_i) (\Delta t_i)^2 \xrightarrow{\mathcal{L}^2} 0.$$

- Con (B), análogo a (II), se desarrolla con la norma $\mathcal{L}^2$:

$$\mathbb{E} \left[\left(\sum_i \frac{\partial^2 V}{\partial x^2} f(t_i) g(t_i) \Delta t_i \Delta W_i \right)^2 \right] = \sum_i \mathbb{E} \left[\left(\frac{\partial^2 V}{\partial x^2} f(t_i) g(t_i) \Delta t_i \right)^2 \right] \Delta t_i.$$

Así que, igual que antes:

$$\sum_i \mathbb{E} \left[\left(\frac{\partial^2 V}{\partial x^2} f(t_i) g(t_i) \Delta t_i \right)^2 \right] \Delta t_i \leq K^2 \sum_i (\Delta t_i)^3 \leq K^2 \max_j (\Delta t_j) \sum_i (\Delta t_i)^2 \rightarrow 0.$$

Por tanto,

$$\sum_i \frac{\partial^2 V}{\partial x^2} f(t_i) g(t_i) \Delta t_i \Delta W_i \xrightarrow{\mathcal{L}^2} 0.$$

- (C) es el núcleo de la fórmula de Itô. Para mostrar la convergencia de esta suma, se va a descomponer (C) restando y sumando $\frac{1}{2} \sum_i \frac{\partial^2 V}{\partial x^2} g^2(t_i) \Delta t_i$, de forma que

$$\sum_i \beta_i (\Delta W_i)^2 = \underbrace{\sum_i \beta_i \Delta t_i}_{\text{(III)}} - \underbrace{\sum_i \beta_i [(\Delta W_i)^2 - \Delta t_i]}_{\text{(IV)}},$$

donde $\beta_i = \frac{1}{2} \frac{\partial^2 V}{\partial x^2} g^2(t_i)$. Para (III) se tiene la convergencia a la integral de Riemann,

$$\frac{1}{2} \sum_i \beta_i \Delta t_i \rightarrow \frac{1}{2} \int_{t_0}^t \frac{\partial^2 V}{\partial x^2}(s, X_s) g^2(s) ds.$$

Para (IV), como los incrementos brownianos son independientes y β_i es $\mathcal{F}_{t_i}$ -medible:

$$\mathbb{E} \left[\left(\sum_i \beta_i [(\Delta W_i)^2 - \Delta t_i] \right)^2 \right] = \sum_i \mathbb{E}[\beta_i^2] \mathbb{E} [((\Delta W_i)^2 - \Delta t_i)^2].$$

Los términos $[(\Delta W_i)^2 - \Delta t_i]$ son independientes y $\mathbb{E}[(\Delta W_i)^2 - \Delta t_i] = 0$. Usando la propiedad $\mathbb{E}[X^4] = 3\sigma^4$ para $X \sim \mathcal{N}(0, \sigma^2)$, se obtiene

$$\begin{aligned} \mathbb{E} [((\Delta W_i)^2 - \Delta t_i)^2] &= \mathbb{E}[(\Delta W_i^4)] + (\Delta t_i)^2 - 2\Delta t_i \mathbb{E}[(\Delta W_i)^2] \\ &= 3(\Delta t_i)^2 + (\Delta t_i)^2 - 2(\Delta t_i)^2 \\ &= 2(\Delta t_i)^2. \end{aligned}$$

Como $\frac{\partial^2 V}{\partial x^2}$ y g son acotadas, $|\beta_i| \leq M$,

$$\sum_i \mathbb{E}[\beta_i^2] 2(\Delta t_i)^2 \leq 2M^2 \max_j (\Delta t_j) \sum_i (\Delta t_i) \rightarrow 0.$$

Así que,

$$\sum_i \beta_i [(\Delta W_i)^2 - \Delta t_i] \xrightarrow{\mathcal{L}^2} 0.$$

Por tanto, esto significa que

$$\frac{1}{2} \sum_i \frac{\partial^2 V}{\partial x^2} g^2(t_i) (\Delta W_i)^2 \rightarrow \frac{1}{2} \int_{t_0}^T \frac{\partial^2 V}{\partial x^2} g^2(t) dt$$

Por último, el resto

$$\sum_i R_i \xrightarrow{\mathcal{L}^2} 0$$

al ser términos de orden menor que Δt_i . En conclusión, juntando todos los anteriores resultados:

$$Y_t - Y_{t_0} = \int_{t_0}^t \left(\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(s) + \frac{1}{2} \frac{\partial^2 V}{\partial x^2} g^2(s) \right) ds + \int_{t_0}^t \frac{\partial V}{\partial x} g(s) dW_s$$

En forma diferencial:

$$dY_t = \left(\frac{\partial V}{\partial t}(t, X_t) + \frac{\partial V}{\partial x}(t, X_t) f(t) + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(t, X_t) g^2(t) \right) dt + \frac{\partial V}{\partial x}(t, X_t) g(t) dW_t.$$

Finalmente, reagrupando:

$$dY_t = \frac{\partial V}{\partial t}(t, X_t) dt + \frac{\partial V}{\partial x}(t, X_t) (f(t) dt + g(t) dW_t) + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(t, X_t) g^2(t),$$

□

La fórmula de Itô se puede generalizar a dimensiones superiores.

Teorema 2.4.3 (Fórmula de Itô multidimensional). *Sea*

$$V : [t_0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^d$$

una función con derivadas parciales continuas $\frac{\partial V}{\partial t}$, $\frac{\partial V}{\partial x_i}$, $\frac{\partial^2 V}{\partial x_i \partial x_j}$ y $\{X_t\}$ un proceso de Itô multidimensional. Entonces, $Y(t) = V(t, X_t)$ satisface

$$\begin{aligned} dY_k(t) = & \left(\frac{\partial V_k}{\partial t}(t, X_t) + \sum_{i=1}^n \frac{\partial V_k}{\partial x_i}(t, X_t) \cdot f_i(t, X_t) \right. \\ & + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 V_k}{\partial x_i \partial x_j}(t, X_t) \cdot \left(\sum_{l=1}^d g_{il}(t, X_t) g_{jl}(t, X_t) \right) \Bigg) dt \\ & + \sum_{i=1}^n \frac{\partial V_k}{\partial x_i}(t, X_t) \cdot \sum_{j=1}^d g_{ij}(t, X_t) dW_j. \end{aligned}$$

De forma más compacta

$$\begin{aligned} dY(t, X_t) = & \left[\frac{\partial V}{\partial t}(t, X_t) + f^T \nabla_x V(t, X_t) + \frac{1}{2} \text{tr} \left(g^T (\nabla_x^2 V(t, X_t)) g \right) \right] dt \\ & + (\nabla_x V(t, X_t))^T g dW_t, \end{aligned}$$

donde $\nabla_x V$ es el gradiente, $\nabla_x^2 V$ es el hessiano de la función $V(t, X_t)$, y $\text{tr}(A) = \sum_{i=1}^n a_{ii}$ es la traza de la matriz $A = (a_{ij}) \in \mathbb{R}^{n \times n}$.

Demostración. Similar al caso unidimensional.

Capítulo 3

Ecuaciones diferenciales estocásticas

3.1 Teorema de existencia y unicidad

Definición 3.1.1 (Ecuación Diferencial Estocástica (de Itô)). Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $\{W_t\}_{t \in T}$ un proceso de Wiener escalar adaptado a la filtración $\{\mathcal{F}_t\}_{t \in T}$. Dadas las funciones $a : T \times \mathbb{R} \rightarrow \mathbb{R}$ y $b : T \times \mathbb{R} \rightarrow \mathbb{R}$ y c una variable aleatoria $\mathcal{F}_{t_0}$ -medible, una ecuación diferencial estocástica (EDE) es una ecuación de la forma:

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \quad X_{t_0} = c, t \in T$$

donde a y b son el coeficiente de deriva y coeficiente de difusión respectivamente.

La SDE anterior representa la ecuación integral:

$$X_t = c + \int_{t_0}^t a(s, X_s)ds + \int_{t_0}^t b(s, X_s)dW_s.$$

Definición 3.1.2. Una solución de la EDE en $[t_0, T]$ es un proceso $\{X_t\}_{t \in [t_0, T]}$ tal que:

1. X_t es $\mathcal{F}_t$ -medible para $t \in [t_0, T]$.
2. $\bar{a}(t, \omega) = a(t, X_t(\omega)) \in \mathcal{L}^1[t_0, T]$ y $\bar{b}(t, \omega) = b(t, X_t(\omega)) \in \mathcal{M}^2[t_0, T]$ c.p.t.
3. La ecuación integral se cumple para todo $t \in [t_0, T]$ c.p.t.

Lema 3.1.3. (Bellman y Gronwall) Si $g \geq 0$ y h es integrable en $[t_0, T]$ y si

$$g(t) \leq L \int_{t_0}^t g(s) ds + h(t), \quad t_0 \leq t \leq T,$$

para $L > 0$, entonces

$$g(t) \leq h(t) + L \int_{t_0}^t e^{L(t-s)} h(s) ds, \quad t_0 \leq t \leq T.$$

Teorema 3.1.4. Considérese la ecuación diferencial estocástica

$$dX_t = a(t, X_t) dt + b(t, X_t) dW_t, \quad X_{t_0} = c \quad t \in [t_0, T]$$

donde $\{W_t\}$ es un proceso de Wiener, c es una variable aleatoria $\mathcal{F}_{t_0}$ -medible, $a : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ y $b : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ son funciones medibles que satisfacen

1. Condición de Lipschitz global: para todo $t \in [t_0, T]$, $x, y \in \mathbb{R}$

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq K|x - y|.$$

2. Crecimiento acotado linealmente: para todo $t \in [t_0, T]$, $x \in \mathbb{R}$

$$|a(t, x)| + |b(t, x)| \leq K(1 + |x|)$$

para cierta constante $K > 0$. Entonces la EDE tiene una solución única $\{X_t\}_{t \in [t_0, T]}$ tal que $X_\omega(t) = X(t)$ es continua con probabilidad 1 y satisface la condición inicial $X_{t_0} = c$.

Demostración.

(a) Unicidad: Supongase que X_t e Y_t son soluciones de la EDE. Entonces:

$$\begin{aligned} \mathbb{E}[|X_t - Y_t|^2] &= \mathbb{E} \left[\left| \int_{t_0}^t (a(s, X_s) - a(s, Y_s)) ds + \int_{t_0}^t (b(s, X_s) - b(s, Y_s)) dW_s \right|^2 \right] \\ &\stackrel{(1)}{\leq} 2\mathbb{E} \left[\left| \int_{t_0}^t (a(s, X_s) - a(s, Y_s)) ds \right|^2 \right] + 2\mathbb{E} \left[\left| \int_{t_0}^t (b(s, X_s) - b(s, Y_s)) dW_s \right|^2 \right] \\ &\stackrel{(2)}{\leq} 2(t - t_0) \int_{t_0}^t \mathbb{E} [|a(s, X_s) - a(s, Y_s)|^2] ds + 2 \int_{t_0}^t \mathbb{E} [|b(s, X_s) - b(s, Y_s)|^2] ds \\ &\stackrel{(3)}{\leq} 2K^2(t - t_0) \int_{t_0}^t \mathbb{E} [|X_s - Y_s|^2] ds + 2K^2 \int_{t_0}^t \mathbb{E} [|X_s - Y_s|^2] ds \\ &= 2K^2(t - t_0 + 1) \int_{t_0}^t \mathbb{E} [|X_s - Y_s|^2] ds \end{aligned}$$

donde en (1) se ha aplicado $|a + b|^2 \leq 2(|a|^2 + |b|^2)$, en (2) la desigualdad de Cauchy-Schwarz y la isometría de Itô, y en (3) la condición de Lipschitz. Se desea concluir que

$$\mathbb{E}[|X_t - Y_t|^2] = 0, \quad t \in [t_0, T]$$

y para ello, se aplica el lema de Bellman y Gronwall con $h(t) = 0$ y $g(t) = \mathbb{E}[|X_t - Y_t|^2]$, con lo que se obtiene

$$\mathbb{E}[|X_t - Y_t|^2] \leq 0$$

luego $X_t = Y_t$ c.p.t para cada $t \in [t_0, T]$.

(b) Existencia: Primero se trata el caso

$$\mathbb{E}[|c|^2] < \infty.$$

Se va a usar un procedimiento de iteración (de Picard-Lindelöf) con $X_t^{(0)} = c$ y, para $n \geq 1$ y $t \in [t_0, T]$,

$$X_t^{(n)} = c + \int_{t_0}^t a(s, X_s^{(n-1)})ds + \int_{t_0}^t b(s, X_s^{(n-1)})dW_s.$$

En virtud de la condición de crecimiento acotado, $X_t^{(0)}$ es no anticipativo y continuo, y con él lo son todos los procesos $X_t^{(n)}$ para $n \geq 1$. El objetivo es mostrar que $\{X_t^{(n)}\}$ converge uniformemente en $[t_0, T]$ a una solución X_t de la ecuación. La suposición $\mathbb{E}[|c|^2] < \infty$ implica

$$\mathbb{E}[|X_t^{(0)}|^2] < \infty$$

y que esto también se cumple para los procesos subsecuentes $X_t^{(n)}$. Con ello, se puede expresar

$$\mathbb{E}[|X_t^{(n+1)} - X_t^{(n)}|^2] \leq L \int_{t_0}^t \mathbb{E}[|X_s^{(n)} - X_s^{(n-1)}|^2]ds$$

con la constante $L = 2(T - t_0 + 1)K^2$. Si se lleva a cabo una iteración sobre esta desigualdad, usando la fórmula de Cauchy, se obtiene

$$\mathbb{E}[|X_t^{(n)} - X_t^{(n-1)}|^2] \leq \frac{L^n}{(n-1)!} \int_{t_0}^t (t-s)^{n-1} \mathbb{E}[|X_s^{(1)} - X_s^{(0)}|^2]ds.$$

Ahora, bajo la suposición de la condición de crecimiento acotado,

$$\begin{aligned}\mathbb{E}[|X_t^{(1)} - X_t^{(0)}|^2] &< 2(T - t_0 + 1)K^2 \int_{t_0}^t (1 + \mathbb{E}[|c|^2])ds \\ &< L(T - t_0)(1 + \mathbb{E}[|c|^2]) = C.\end{aligned}$$

Por lo que se tiene que

$$\mathbb{E}[|X_t^{(n+1)} - X_t^{(n)}|^2] < C \frac{(L(T - t_0))^n}{n!}, \quad n \geq 0.$$

Para probar la convergencia uniforme de $X_t^{(n)}$ en sí, necesitamos encontrar una estimación para

$$d_n = \sup_{t \in [t_0, T]} |X_t^{(n+1)} - X_t^{(n)}|.$$

Se tiene que

$$\begin{aligned}d_n &\leq \int_{t_0}^t |a(s, X_s^{(n)}) - a(s, X_s^{(n-1)})| ds \\ &\quad + \left| \int_{t_0}^t (b(s, X_s^{(n)}) - b(s, X_s^{(n-1)})) dW_s \right|.\end{aligned}$$

Sabiendo que

$$\mathbb{E} \left[\sup_{a \leq t \leq b} |X_t - X_a|^2 \right] \leq 4 \int_a^b \mathbb{E}[b(s)]^2 ds$$

y aplicando la condición de Lipschitz, se obtiene

$$\begin{aligned}E[d_n^2] &< 2(T - t_0)K^2 \int_{t_0}^T \mathbb{E}[|X_s^{(n)} - X_s^{(n-1)}|^2] ds \\ &\quad + 8K^2 \int_{t_0}^T \mathbb{E}[|X_s^{(n)} - X_s^{(n-1)}|^2] ds.\end{aligned}$$

Por tanto, para $n \geq 0$,

$$\mathbb{E}[d_n^2] \leq (2(T - t_0) + 8)K^2(T - t_0)C \frac{(L(T - t_0))^{n-1}}{(n-1)!} = C_1 \frac{(L(T - t_0))^{n-1}}{(n-1)!}.$$

En virtud del lema de Borel-Cantelli y el criterio de convergencia de Weierstrass, se obtiene la convergencia de la serie

$$\sum_{n=1}^{\infty} \mathbb{P}[d_n > n^{-2}] < C_1 \sum_{n=1}^{\infty} \frac{(L(T - t_0))^{n-1} n^4}{(n-1)!} < \infty.$$

Luego $X_t^{(n)}$ converge uniformemente en $[t_0, T]$ casi seguro a un límite X_t que es la solución:

$$\lim_{N \rightarrow \infty} X_t^{(0)} + \sum_{i=1}^N (X_t^{(i)} - X_t^{(i-1)}) = \lim_{N \rightarrow \infty} X_t^{(N)} = X_t,$$

El caso de una condición inicial general c se reduce al caso $\mathbb{E}[|c|^2] < \infty$ imponiendo

$$c_N = \begin{cases} c, & \text{si } |c| \leq N \\ 0, & \text{en caso contrario} \end{cases}$$

□

3.2 Ecuaciones diferenciales estocásticas lineales

Definición 3.2.1. *Una ecuación diferencial estocástica*

$$dX_t = a(t, X_t) dt + b(t, X_t) dW_t$$

es lineal si $a(t, x)$ y $b(t, x)$ son funciones lineales de $x \in \mathbb{R}^d$ en $[t_0, T] \times \mathbb{R}^d$, es decir, si es de la forma

$$dX_t = (A(t)X_t + a(t)) dt + \sum_{i=1}^n (B_i(t)X_t + b_i(t)) dW_t^i$$

donde $W_t = (W_t^1, \dots, W_t^n)$, $a, b_i : [t_0, T] \rightarrow \mathbb{R}^d$ y $A, B_i : [t_0, T] \rightarrow \mathbb{R}^{d \times d}$, $1 \leq i \leq n$. Además, una ecuación diferencial lineal es homogénea si $a(t) = b_i(t) = 0$, $1 \leq i \leq n$.

El siguiente teorema es una consecuencia inmediata del teorema de existencia y unicidad.

Teorema 3.2.2. *La ecuación diferencial estocástica lineal*

$$dX_t = (A(t)X_t + a(t)) dt + \sum_{i=1}^n (B_i(t)X_t + b_i(t)) dW_t^i,$$

con $A(t)$, $a(t)$, $B_i(t)$ y $b_i(t)$ medibles y acotadas en $[t_0, T]$ y condición inicial $X_{t_0} = c$ independiente de $W_t - W_{t_0}$ ($t \geq t_0$) tiene solución única en $[t_0, T]$. Si las condiciones se cumplen en cada subintervalo de $[t_0, \infty)$ existe una solución global única.

Corolario 3.2.3. *Para ecuación diferencial estocástica lineal autónoma (A, a, B_i y b_i indepen-*

dientes de t)

$$dX_t = (AX_t + a) dt + \sum_{i=1}^n (B_i X_t + b_i) dW_t^i$$

siempre existe solución global única.

Es interesante estudiar explícitamente dicha solución. Para ello se debe conocer el siguiente concepto.

Definición 3.2.4. La solución fundamental Φ de la ecuación estocástica lineal homogénea

$$dX_t = A(t)X_t dt + \sum_{i=1}^n B_i(t)X_t dW_t^i$$

sobre $[t_0, T]$ es el proceso estocástico

$$\Phi : [t_0, T] \times \Omega \longrightarrow \mathbb{R}^{d \times d}$$

de forma que el elemento i -ésimo de Φ es la solución de dicha EDE con $X_{t_0} = (0, \dots, 0, \overset{i}{1}, 0, \dots)$.

El siguiente teorema ofrece una expresión explícita para la solución fundamental EDE homogénea en el caso particular unidimensional.

Teorema 3.2.5. La solución fundamental Φ_t de la EDE lineal homogénea unidimensional es

$$\Phi_t = \exp \left\{ \int_{t_0}^t \left(A(s) - \frac{1}{2} \sum_{i=1}^n B_i(s)^2 \right) ds + \sum_{i=1}^n \int_{t_0}^t B_i(s) dW_s^i \right\}.$$

Demostración. La solución fundamental es la solución de la EDE lineal homogénea con la condición inicial $\Phi_{t_0} = 1$. Considerando

$$Y_t = \int_{t_0}^t \left(A(s) - \frac{1}{2} \sum_{i=1}^n B_i(s)^2 \right) ds + \sum_{i=1}^n \int_{t_0}^t B_i(s) dW_s^i$$

y aplicando la fórmula de Itô con $\Phi_t = e^{Y_t}$

$$\begin{aligned} d\Phi_t &= Y_t dY_t + \frac{1}{2} Y_t \sum_{i=1}^n B_i(t)^2 dt \\ &= Y_t \left(\left(A(t) - \frac{1}{2} \sum_{i=1}^n B_i(t)^2 \right) dt + \sum_{i=1}^n B_i(t) dW_t^i \right) + \frac{1}{2} Y_t \sum_{i=1}^n B_i(t)^2 dt \\ &= Y_t \left(A(t) + \sum_{i=1}^n B_i(t) dW_t^i \right). \end{aligned}$$

Usando la solución fundamental Φ , se puede obtener la solución general de la EDE lineal.

Teorema 3.2.6. *La ecuación diferencial estocástica lineal*

$$dX_t = (A(t)X_t + a(t)) dt + \sum_{i=1}^n (B_i(t)X_t + b_i(t)) dW_t^i, \quad X_{t_0} = c$$

tiene en $[t_0, T]$ solución

$$X_t = \Phi_t \left(c + \int_{t_0}^t \Phi_s^{-1} \left(a(s) - \sum_{i=1}^n B_i(s)b_i(s) \right) ds + \sum_{i=1}^n \int_{t_0}^t \Phi_s^{-1} b_i(s) dW_s^i \right),$$

donde la matriz $\Phi_t \in \mathbb{R}^{d \times d}$ es la solución fundamental de la ecuación homogénea asociada, es decir,

$$d\Phi_t = A(t)\Phi_t dt + \sum_{i=1}^m B_i(t)\Phi_t dW_t^i, \quad \Phi_{t_0} = I_d.$$

Demostración. Se conoce que existe una solución única de la ecuación, luego solo se necesita verificar que X_t satisface la ecuación. Considerando

$$Z_t = c + \int_{t_0}^t \Phi_s^{-1} dY_s,$$

se tiene que

$$dZ_t = \Phi_t^{-1} dY_t$$

donde

$$dY_s = \left(a(s) - \sum_{i=1}^n B_i(s)b_i(s) \right) ds + \sum_{i=1}^n b_i(s) dW_s^i.$$

Usando la fórmula de Itô, aplicada al proceso $X_t = \Phi_t Z_t$ para cada una de sus componentes, se obtiene

$$\begin{aligned} dX_t &= \Phi_t dZ_t + (d\Phi_t)Z_t + \left(\sum_{i=1}^n B_i(t)\Phi_t \Phi_t^{-1} b_i(t) \right) dt \\ &= \Phi_t \Phi_t^{-1} dY_t + A(t)\Phi_t \Phi_t^{-1} X_t dt + \sum_{i=1}^m B_i(t)\Phi_t \Phi_t^{-1} X_t dW_t^i + \left(\sum_{i=1}^n B_i(t)b_i(t) \right) dt \\ &= (a(t) + A(t)X_t) dt + \sum_{i=1}^n (B_i(t)X_t + b_i(t)) dW_t^i. \end{aligned}$$

Con la condición inicial $X_{t_0} = \Phi_{t_0} Z_{t_0} = I c = c$. □

Ahora, se van a presentar algunos de los ejemplos de EDEs lineales más conocidos y usados en la modelización.

Ejemplo 3.2.7 Movimiento browniano aritmético

$$dX_t = a dt + \sigma dW_t$$

La solución es

$$X_t = X_0 + \int_0^t a ds + \int_0^t \sigma dW_s = X_0 + at + \sigma W_t$$

Se verifica que

$$\begin{cases} \mathbb{E}[X_t] = \mathbb{E}[X_0] + at \\ \text{Var}[X_t] = \sigma^2 t \end{cases}$$

Ejemplo 3.2.8 Movimiento browniano geométrico

$$dX_t = aX_t dt + \sigma X_t dW_t$$

Aplicando la fórmula de Itô sobre $Y_t = \ln(X_t)$

$$d(\ln(X_t)) = \left(a - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t$$

con lo que

$$\ln(X_t) = \ln(X_0) + \left(a - \frac{1}{2}\sigma^2\right)t + \sigma W_t.$$

De aquí se obtiene la solución

$$X_t = X_0 e^{(a - \frac{1}{2}\sigma^2)t + \sigma W_t}.$$

Tomando esperanzas:

$$\mathbb{E}[X_t] = \mathbb{E}[X_0] + a \int_0^t \mathbb{E}[X_s] ds.$$

Denotando $m(t) = \mathbb{E}[X_t]$, se tiene que $m(t)$ es la solución de $m'(t) = a m(t)$, con $m(0) = \mathbb{E}[X_0]$.

Entonces

$$m(t) = \mathbb{E}[X_t] = \mathbb{E}[X_0]e^{at}.$$

Ejemplo 3.3.9 Proceso de Ornstein-Uhlenbeck

$$dX_t = -\lambda X_t dt + \sigma dW_t \quad \lambda, \sigma > 0.$$

Se considera el proceso $Y_t = X_t e^{\lambda t}$. Aplicando la fórmula de Itô:

$$dY_t = e^{\lambda t} \sigma dW_t,$$

luego,

$$Y_t = Y_0 + \sigma \int_0^t e^{\lambda s} dW_s.$$

Entonces

$$X_t = e^{-\lambda t} Y_t = e^{-\lambda t} X_0 + e^{-\lambda t} \sigma \int_0^t e^{\lambda s} dW_s,$$

que verifica

$$\begin{cases} \mathbb{E}[X_t] = e^{-\lambda t} Y_0 \\ \text{Var}[X_t] = \frac{\sigma^2}{2\lambda} (1 - e^{-2\lambda t}) \end{cases}$$

3.3 Estabilidad de las soluciones

Considérese la EDE autónoma:

$$dX_t = a(X_t)dt + b(X_t)dW_t, \quad t \in [t_0, T]$$

donde W_t es un proceso de Wiener y $a, b : T \times \mathbb{R} \rightarrow \mathbb{R}$ son funciones medibles que satisfacen condiciones de Lipschitz global y crecimiento lineal.

Definición 3.3.1. *Un proceso estocástico X_t es una solución de equilibrio para la EDE si $X_t \equiv c$ (constante) casi seguramente para todo $t \in [t_0, T]$ y satisface $a(c) = 0$ y $b(c) = 0$. En particular, $c = 0$ define la solución de equilibrio trivial.*

Definición 3.3.2. *La solución de equilibrio $X_t = 0$ es estable en media cuadrática si para todo $\epsilon > 0$, existe $\delta > 0$ tal que $\mathbb{E}[X_{t_0}^2] < \delta$ implica que $\mathbb{E}[X_t^2] < \epsilon$ para todo $t \geq t_0$, y además $\lim_{t \rightarrow \infty} \mathbb{E}[X_t^2] = 0$.*

Considérese la EDE lineal homogénea (movimiento browniano geométrico):

$$dX_t = \lambda X_t dt + \mu X_t dW_t, \quad X_0 = 1.$$

Como ya se ha visto, su solución es

$$X_t = e^{(a - \frac{1}{2}\sigma^2)t + \sigma W_t}.$$

Tomando esperanza cuadrática:

$$\mathbb{E}[X_t^2] = \mathbb{E}[e^{2(\lambda - \frac{1}{2}\sigma^2)t + 2\sigma W_t}]$$

Con las propiedades del proceso de Wiener y la distribución normal, se tiene que

$$2(\lambda - \frac{1}{2}\sigma^2)t + 2\sigma W_t \sim \mathcal{N}\left(2(\lambda - \frac{1}{2}\sigma^2)t, 4\sigma^2 t\right)$$

luego resulta que

$$\mathbb{E}[X_t^2] = e^{(2\lambda + \mu^2)t}.$$

La estabilidad en media cuadrática, $\mathbb{E}[X_t^2] \rightarrow 0$, equivale a:

$$\lim_{t \rightarrow \infty} e^{(2\lambda + \mu^2)t} = 0 \iff 2\lambda + \mu^2 < 0.$$

Se denomina región de estabilidad de la ecuación lineal a $\mathcal{R} = \{(\lambda, \mu) : 2\lambda + \mu^2 < 0\}$, pudiéndose representar en el plano (λ, μ^2) .

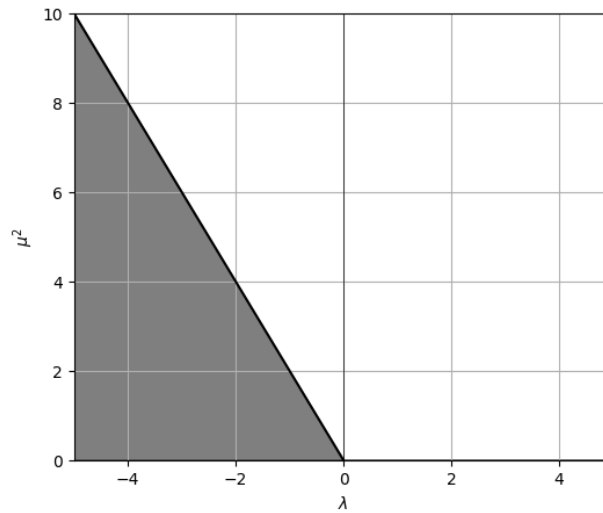


Figura 3.3.1: Región de estabilidad en media cuadrática en el plano (λ, μ^2) .

Capítulo 4

Métodos numéricos en EDEs

En este capítulo se van definir y analizar algunos de los métodos numéricos más utilizados en la resolución de ecuaciones diferenciales estocásticas.

Durante esta sección, se supondrá que la EDE cuya solución se va aproximar es de la forma

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \quad X_{t_0} = c, \quad t \in [t_0, T]$$

y que verifica las hipótesis del teorema de existencia y unicidad. Su solución será denotada X . Un método numérico es un algoritmo que proporciona un proceso discreto que aproxima a la solución exacta.

Se considera una discretización temporal del intervalo $[t_0, T]$: $t_0 < t_1 < \dots < t_N = T$, $N \in \mathbb{N}$ y paso que, por simplicidad, se asume constante $h = (T - t_0)/N$. Cada método consiste en definir

$$Y^h : P^h \times \Omega \rightarrow \mathbb{R}^n$$

con $P^h = \{t_0, t_0 + h, \dots, T - h, T\}$, que va a aproximar los valores de X a cada punto de P^h . Se denota a la aproximación numérica $Y^h(t_0 + nh)$ como $Y_{t_n}^h$, o solo Y_n si h es claramente definido.

4.1 Convergencia de los métodos numéricos

La convergencia del método numérico mide la calidad de la aproximación cuando el tamaño de paso tiende a 0. Para un método numérico estocástico se diferencia entre convergencia fuerte y convergencia débil.

Definición 4.1.1 (Convergencia débil). *Se dice que la aproximación discreta $\{Y_0, Y_1, \dots, Y_n\}$, con un paso h , converge débilmente a la solución X con orden β si para cada $g \in \mathcal{C}^{2\beta+2}$ existe una constante $K > 0$ independiente de h tal que*

$$|\mathbb{E}[g(Y_N) - g(X_T)]| \leq K_g h^\beta$$

Definición 4.1.2 (Convergencia fuerte). *Una aproximación discreta $\{Y_0, Y_1, \dots, Y_n\}$, con un paso h , converge fuertemente a la solución X con orden γ en el tiempo T si existe una constante $C > 0$ independiente de h y un $\delta_0 > 0$ tal que para cada $h \in (0, \delta_0)$*

$$\mathbb{E}[X_T - Y_N] \leq Ch^\gamma$$

La convergencia fuerte se centra en medir la diferencia entre las trayectorias exacta y aproximada (en media cuadrática), mientras que la convergencia débil mide la aproximación de las distribuciones o sus momentos.

4.2 Método de Euler-Maruyama

El método de Euler-Maruyama es la extensión natural del método de Euler determinista. Se obtiene integrando la EDE

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \quad X_{t_0} = c, \quad t \in [t_0, T]$$

sobre $[t_n, t_n + 1]$ y aproximando los integrandos por su valor a la izquierda:

$$\begin{aligned} X_{n+1} &= X_n + \int_{t_n}^{t_{n+1}} a(s, X_s) ds + \int_{t_n}^{t_{n+1}} b(s, X_s) dW_s \\ &\approx X_n + a(t_n, X_n)(t_{n+1} - t_n) + b(t_n, X_n)(W_{t_{n+1}} - W_{t_n}) \end{aligned}$$

Entonces se tiene el esquema

$$\begin{cases} Y_0 = c \\ Y_{n+1} = Y_n + a(t_n, Y_n)h + b(t_n, Y_n)\Delta W_n, \quad n = 0, 1, \dots, N-1 \end{cases}$$

donde $\Delta W_n = W_{t_{n+1}} - W_{t_n} \sim \mathcal{N}(0, h)$ son incrementos Brownianos independientes.

Teorema 4.2.1. *Bajo las condiciones de Lipschitz y crecimiento lineal, el método de Euler-Maruyama tiene orden de convergencia fuerte $\gamma = 0.5$. Es decir, existe una constante $C > 0$ independiente de h tal que:*

$$\mathbb{E}[|X_T - Y_N|] \leq Ch^{1/2}$$

Teorema 4.2.2. *Bajo las condiciones de Lipschitz, crecimiento lineal, y asumiendo que $a(x)$, $b(x)$ son suficientemente diferenciables con derivadas acotadas, el método de Euler-Maruyama tiene orden de convergencia débil $\beta = 1$. Es decir, para cada función polinómica $g(x)$ existe una constante $K_g > 0$ independiente de h tal que:*

$$|\mathbb{E}[g(X_T)] - \mathbb{E}[g(Y_N)]| \leq K_g h$$

4.3 Método de Milstein

El método de Milstein, desarrollado por Grigori Milstein, es un método numérico que mejora la aproximación del método de Euler-Maruyama. Mientras que el método de Euler-Maruyama solo captura el término de difusión de forma aproximada, el método de Milstein mejora la precisión incorporando la derivada del coeficiente de difusión, capturando parte de la correlación entre el ruido y su amplitud. Se obtiene aplicando la fórmula de Itô al coeficiente

de difusión en la integral estocástica:

$$\begin{aligned}\int_{t_n}^{t_{n+1}} b(s, Y_s) dW_s &= \int_{t_n}^{t_{n+1}} \left(b(X_n) + b'(X_n) b(X_n) \int_{t_n}^s dW_u + \dots \right) dW_s \\ &= b(X_n) \Delta W_n + b'(X_n) b(X_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^s dW_u dW_s + R\end{aligned}$$

donde el término residual R es de orden superior. La integral doble iterada se calcula:

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^s dW_u dW_s = \int_{t_n}^{t_{n+1}} (W_s - W_{t_n}) dW_s = \frac{1}{2} [(\Delta W_n)^2 - h]$$

usando la fórmula de Itô para $d(W_t^2) = 2W_t dW_t + dt$.

Entonces, resulta el esquema

$$\begin{cases} Y_0 = c \\ Y_{n+1} = Y_n + a(Y_n)h + b(Y_n)\Delta W_n + \frac{1}{2}b'(Y_n)b(Y_n)[(\Delta W_n)^2 - h] \end{cases}$$

donde $n = 0, 1, \dots, N-1$, y $\Delta W_n = W_{t_{n+1}} - W_{t_n} \sim \mathcal{N}(0, h)$ son incrementos Brownianos independientes.

Teorema 4.3.1. *Bajo las hipótesis de Lipschitz y crecimiento lineal para $a(x)$, $b(x)$, y asumiendo además que $b'(x)$ es Lipschitz y de crecimiento lineal, el método de Milstein tiene orden de convergencia fuerte y débil igual a 1.*

4.4 Aproximaciones de Itô-Taylor

En 1982, Platen y Wagner propusieron una generalización estocástica de la fórmula de Taylor para procesos de Itô de tipo difusión. La generalización se basó en el uso de la regla de la cadena estocástica (fórmula de Itô) y se expresa en términos de integrales estocásticas múltiples. Las expansiones de Itô-Taylor generalizan los métodos de Euler-Maruyama y de Milstein. Permiten construir esquemas de cualquier orden de convergencia (fuerte o débil) truncando una expansión estocástica de Taylor de la solución. Antes de todo, se van a presentar distintas definiciones que van a permitir simplificar la notación y entender mejor las expansiones de Itô-Taylor.

Se va considerar el multi-índice $\alpha = (j_1, \dots, j_l)$, $j_k \in \{0, 1\}$, de longitud $l = l(\alpha) \in \{1, 2, \dots\}$. A su vez, se define la función n que va a contabilizar cuantos componentes de α son iguales a 0, escrito por $n(\alpha)$. Se denota $\mathcal{M}$ al conjunto de todos los multi-índices

$$\mathcal{M} = \{(j_1, \dots, j_l) : j_k \in \{0, 1\}, k \in \{1, \dots, l\}, l = 1, 2, \dots\} \cup \{v\}$$

donde v es, por convenio, el multi-índice de longitud 0. Si $\alpha = (j_1, \dots, j_l) \in \mathcal{M}$ se define $-\alpha = (j_2, j_3, \dots, j_l)$. Dado un subconjunto $\mathcal{A} \in \mathcal{M}$, el conjunto restante $\mathcal{R}(\mathcal{A})$ se define:

$$\mathcal{R}(\mathcal{A}) = \{\alpha \in \mathcal{M} \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}$$

Se dice que un subconjunto finito $\mathcal{A}$ de $\mathcal{M}$ es jerárquico si para cada $\alpha \in \mathcal{A}$ se tiene que $-\alpha \in \mathcal{A}$. En particular, si $n \in \mathbb{N}$ el conjunto de multi-índices de longitud menor o igual que n

$$\mathcal{M}_n = \{\alpha \in \mathcal{M} : l(\alpha) \leq n\}$$

es un conjunto jerárquico.

De la fórmula de Itô unidimensional, se definen los operadores asociados a la ecuación diferencial estocástica de coeficientes a, b por:

$$\begin{aligned} L^{(0)} &= \frac{\partial}{\partial t} + a(t, x) \frac{\partial}{\partial x} + \frac{1}{2} b^2(t, x) \frac{\partial^2}{\partial x^2} \\ L^{(1)} &= b(t, x) \frac{\partial}{\partial x} \end{aligned}$$

Tomando multi-índices $\alpha = (j_1, \dots, j_l)$, se definen los operadores compuestos

$$L^\alpha = L^{j_1} \circ L^{j_2} \circ \dots \circ L^{j_l}.$$

Por otro lado, para simplificar, se renombra las integrales de Itô dW_t como dW_t^1 y las integrales de Lebesgue dt como dW_t^0 . Con esto, las integrales estocásticas múltiples se representan como integrales respecto a un multi-índice de la forma

$$\int_{t_0}^t f(\cdot, X) d\alpha = \int_{t_0}^t \dots \int_{t_0}^{s_3} \int_{t_0}^{s_2} f(s_1, X_{s_1}) dW_{s_1}^{j_1} dW_{s_2}^{j_2} \dots dW_{s_l}^{j_l}.$$

En particular para $f \equiv 1$ se escribe $I_\alpha = \int_{t_0}^t d\alpha$.

Entonces, si

$$X_t = X_{t_0} + \int_{t_0}^t f(s) ds + \int_{t_0}^t g(s) dW_s$$

es un proceso de Itô unidimensional con respecto a un proceso de Wiener escalar y $\mathcal{A} \in \mathcal{M}$ un conjunto jerárquico de multi-índices, la expansión de Itô-Taylor de $f : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ con respecto a $\mathcal{A}$ es

$$f(t, X_t) = \sum_{\alpha \in \mathcal{A}} (L^\alpha f)(t_0, X_{t_0}) I_\alpha + \sum_{\alpha \in R(\mathcal{A})} \int_{t_0}^t (L^\alpha f)(\cdot, X_\cdot) d\alpha$$

Nótese que la definición de la expansión de Itô-Taylor se caracteriza por la elección del conjunto jerárquico $\mathcal{A}$, luego se tienen algunos casos particulares.

Con $n \in \mathbb{N}$, la expansión de Itô-Taylor asociada al conjunto jerárquico

$$\mathcal{M}_n = \{\alpha \in \mathcal{M} : l(\alpha) \leq n\}$$

es la expansión de Itô-Taylor débil de orden n :

$$f(t, X_t) = \sum_{\alpha \in \mathcal{M}_n} (L^\alpha f)(t_0, X_{t_0}) I_\alpha + \sum_{\alpha \in R(\mathcal{M}_n)} \int_{t_0}^t (L^\alpha f)(\cdot, X_\cdot) d\alpha$$

Despreciando el resto, se obtiene la expansión de Itô-Taylor débil truncada de orden n :

$$f(t, X_t) \approx \sum_{\alpha \in \mathcal{M}_n} (L^\alpha f)(t_0, X_{t_0}) I_\alpha$$

Cuando $f(t, x) = x$ se obtiene la expansión de Itô-Taylor truncada de orden n de la solución X_t :

$$\eta_n(t) = \sum_{\alpha \in \mathcal{M}_n} (L^\alpha f)(t_0, X_{t_0}) I_\alpha$$

Bajo condiciones suficientes sobre los coeficientes de la ecuación, $\eta_n(t)$ es una aproximación de la solución exacta X_t de orden local $n + 1$.

Se denota C_P^m al espacio de funciones m veces diferenciables con continuidad tales que todas sus derivadas parciales de orden hasta e incluyendo m tienen crecimiento polinomial.

Teorema 4.4.1. *Sea $n \in \mathbb{N}$ y $\Delta \geq 0$. Sea X_t la solución de la EDE con coeficientes $a, b \in C_P^{2n+2}$ junto con la condición inicial $X_0 = x_0$. Si $\eta_n(t)$ es la expansión truncada de*

orden n de X_t entonces para cada $g \in C_P^{2n+2}$ existe una constante $K > 0$ tal que

$$\sup_{0 \leq t \leq \Delta} |E[g(X_t) - g(\eta_n(t))]| \leq K \Delta^{n+1}$$

Por otro lado, si $\gamma = 0.5, 1, 1.5, \dots$, entonces la expansión de Itô-Taylor asociada al conjunto jerárquico

$$A_\gamma = \left\{ \alpha \in \mathcal{M} : l(\alpha) + n(\alpha) \leq 2\gamma \text{ ó } n(\alpha) = \gamma + \frac{1}{2} \right\}$$

es la expansión de Itô-Taylor fuerte de orden γ :

$$f(t, X_t) = \sum_{\alpha \in A_\gamma} (L^\alpha f)(t_0, X_{t_0}) I_\alpha + \sum_{\alpha \in R(A_\gamma)} \int_{t_0}^t (L^\alpha f)(\cdot, X_\cdot) d\alpha$$

Despreciando el resto obtenemos la expansión de Itô-Taylor fuerte truncada de orden γ :

$$f(t, X_t) \approx \sum_{\alpha \in A_\gamma} (L^\alpha f)(t_0, X_{t_0}) I_\alpha$$

Cuando $f(t, x) = x$, se obtiene la expansión de Itô-Taylor truncada de orden fuerte γ de la solución X_t :

$$Y_\gamma(t) = \sum_{\alpha \in A_\gamma} (L^\alpha f)(t_0, X_{t_0}) I_\alpha$$

El esquema obtenido con la expansión de Itô-Taylor truncada $Y_\gamma(t)$ tiene orden fuerte γ .

Teorema 4.4.2. *Bajo las condiciones de Lipschitz y crecimiento lineal, un esquema basado en una expansión de Itô-Taylor truncada que incluye todos los términos cuyas integrales iteradas asociadas tienen multi-índice α cumpliendo $l(\alpha) + n(\alpha) \leq 2\gamma$ ó $n(\alpha) = \gamma + \frac{1}{2}$, tiene orden de convergencia fuerte γ .*

Para EDEs multidimensionales respecto a procesos de Wiener m -dimensionales, la expansión de Itô-Taylor se generaliza utilizando la fórmula de Itô multidimensional y permitiendo a los multi-índices $\alpha = (j_1, \dots, j_l)$ que $j_k \in \{0, 1, \dots, m\}$.

4.5 Métodos implícitos

Los métodos explícitos (Euler-Maruyama, Milstein) pueden volverse inestables para problemas rígidos, requiriendo pasos de tiempo muy pequeños. Los métodos implícitos mejoran la estabilidad mediante la introducción de implicitud controlada en el término determinista.

Para la EDE general, el esquema de los métodos implícitos de Euler o métodos θ se definen:

$$\begin{cases} Y_0 = c \\ Y_{n+1} = Y_n + ((1 - \theta)a(t_n, Y_n) + \theta a(t_{n+1}, Y_{n+1}))h + b(t_n, Y_n)\Delta W_n \end{cases}$$

donde θ es el parámetro de implicitud. El término estocástico típicamente permanece explícito por dificultades técnicas. Existen algunos casos particulares:

- $\theta = 0$: método de Euler-Maruyama.
- $\theta = 0.5$: regla del trapecio.
- $\theta = 1$: método de Euler-Maruyama implícito.

Teorema 4.5.1. *Bajo las condiciones de Lipschitz y crecimiento lineal, para $\theta \in [\frac{1}{2}, 1]$, el método tiene orden de convergencia fuerte $\gamma = 0.5$:*

$$\mathbb{E}[|X_T - Y_N|] \leq Ch^{1/2}.$$

Teorema 4.5.2. *Bajo las condiciones de Lipschitz y crecimiento lineal, para $\theta \in [0, 1]$, el método tiene orden de convergencia débil $\beta = 1$:*

$$|\mathbb{E}[g(X_T)] - \mathbb{E}[g(Y_N)]| \leq C_g h.$$

La ventaja principal de los métodos implícitos no es el orden de convergencia, sino la estabilidad para pasos grandes. Como ya se comentó, la estabilidad (lineal) en media cuadrática se basa en el estudio de la función test (EDE lineal homogénea). Para dicha ecuación, la condición de estabilidad se reduce a

$$\lim_{t \rightarrow \infty} \mathbb{E}[|X_t|^2] = 0 \iff 2\lambda + \mu^2 < 0.$$

El esquema de Euler implícito aplicado a la función test es:

$$Y_{n+1} = Y_n + \theta \lambda Y_{n+1} h + (1 - \theta) \lambda Y_n h + \mu Y_n \Delta W_n.$$

Despejando Y_{n+1} :

$$Y_{n+1} = Y_n \frac{1 + (1 - \theta) \lambda h + \mu \Delta W_n}{1 - \theta \lambda h}.$$

La estabilidad en media cuadrática de la solución numérica se reduce a

$$\lim_{n \rightarrow \infty} \mathbb{E}[|Y_n|^2] = 0.$$

Como Y_n es $\mathcal{F}_{t_n}$ -medible y ΔW_n es independiente de $\mathcal{F}_{t_n}$, tomando esperanzas

$$\begin{aligned} \mathbb{E}[|Y_{n+1}|^2] &= |Y_n|^2 \mathbb{E} \left[\left| \frac{1 + (1 - \theta) \lambda h + \mu \Delta W_n}{1 - \theta \lambda h} \right|^2 \right] \\ &= |Y_n|^2 \cdot \frac{1}{(1 - \theta \lambda h)^2} \cdot \mathbb{E}_n \left[(1 + (1 - \theta) \lambda h + \mu \Delta W_n)^2 \right] \end{aligned}$$

Usando propiedades de $\Delta W_n \sim \mathcal{N}(0, h)$ y $\mathbb{E}[(\Delta W_n)^2] = h$:

$$\begin{aligned} \mathbb{E} \left[|1 + (1 - \theta) \lambda h + \mu \Delta W_n|^2 \right] &= \mathbb{E} \left[(1 + (1 - \theta) \lambda h)^2 + 2(1 + (1 - \theta) \lambda h) \mu \Delta W_n \right. \\ &\quad \left. + \mu^2 (\Delta W_n)^2 \right] = (1 + (1 - \theta) \lambda h)^2 + \mu^2 h \end{aligned}$$

Entonces,

$$\mathbb{E}[|Y_{n+1}|^2] = |Y_n|^2 \frac{(1 + (1 - \theta) \lambda h)^2 + \mu^2 h}{(1 - \theta \lambda h)^2}$$

Por tanto, la solución numérica es estable en media cuadrática si y sólo si:

$$\frac{(1 + (1 - \theta) \lambda h)^2 + \mu^2 h}{(1 - \theta \lambda h)^2} < 1.$$

Desarrollando queda

$$(1 - 2\theta) \lambda^2 h + (2\lambda + \mu^2) < 0,$$

donde $h > 0$ es el tamaño del paso. Se observa que para $\theta \geq 0.5$ la suma siempre es negativa, luego se tiene una estabilidad incondicional. Para $\theta < 0.5$, se tiene estabilidad si

$$h < \frac{2\lambda + \mu^2}{(1 - 2\theta) \lambda^2}.$$

En cuanto a la interpretación geométrica, se puede parametrizar la región de estabilidad en el plano $x = \lambda h$ e $y = \mu^2 h$ para el parámetro θ , resultando en:

$$\mathcal{R}(\theta) = \{(x, y) : (1 - 2\theta)x^2 + 2x + y < 0\}$$

Si se toma $\theta = 0.5$ y $h = 1$ se obtiene

$$\mathcal{R}(0.5) = \{(x, y) : 2x + y < 0\}$$

que es justamente la región de estabilidad de la función test lineal. Para $\theta > 0.5$, se observa que la región de estabilidad contiene a la región de estabilidad de la función test, lo que indica el beneficio de los métodos implícitos. Esta idea se muestra en la figura 4.5.1, donde $\mathcal{R}(\theta)$ es el área roja y la región de estabilidad lineal es el área azul.

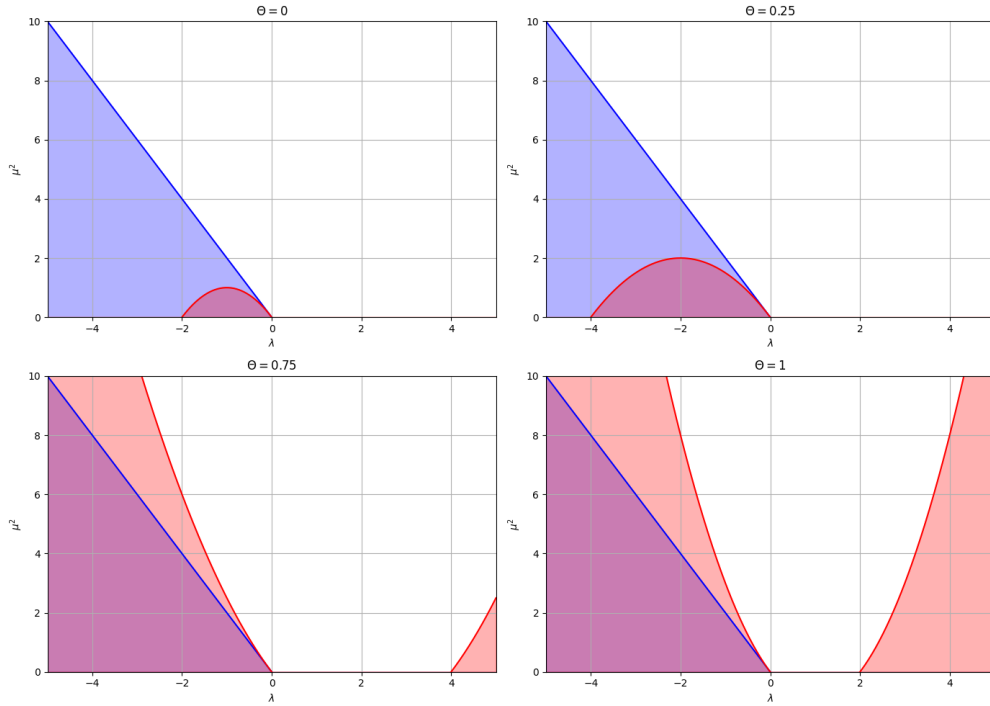


Figura 4.5.1: Regiones de estabilidad de los métodos θ de Euler en el plano (λ, μ^2) .

También se desarrolla el método implícito θ -Milstein

$$\begin{cases} Y_0 = c \\ Y_{n+1} = Y_n + (\theta a(t_{n+1}, Y_{n+1})) + (1 - \theta)a(t_n, Y_n)h + b(t_n, Y_n)\Delta W_n \\ \quad + \frac{1}{2}b(t_n, Y_n)\frac{\partial b}{\partial x}(t_n, Y_n)(\Delta W_n^2 - h) \end{cases}$$

Se conoce que este esquema tiene orden de convergencia fuerte $\gamma = 1$ bajo las condiciones de suficientes de existencia y unicidad y si $\frac{\partial b}{\partial x}$ es Lipschitz.

4.6 Métodos predictor-corrector

Los métodos predictor-corrector combinan la estabilidad de métodos implícitos con la facilidad de implementación de explícitos, evitando resolver ecuaciones no lineales explícitas. La estructura general es primero calcular la aproximación preliminar $\tilde{Y}_{n+1}$ (paso predictor o explícito) y luego mejorar la aproximación usando $\tilde{Y}_{n+1}$ en los términos implícitos (paso corrector o implícito). El esquema predictor-corrector más popular se obtiene con los métodos de Euler implícitos:

$$\left\{ \begin{array}{l} Y_0 = c \\ \text{Predictor (Euler-Maruyama):} \\ \quad \tilde{Y}_{n+1} = Y_n + a(t_n, Y_n)h + b(t_n, Y_n)\Delta W_n \\ \text{Corrector } (\theta\text{-Implícito):} \\ \quad Y_{n+1} = Y_n + \left[(1 - \theta)a(t_n, Y_n) + \theta a(t_{n+1}, \tilde{Y}_{n+1}) \right] h + b(t_n, Y_n)\Delta W_n \end{array} \right.$$

Se puede demostrar que, bajo las condiciones de Lipschitz y crecimiento lineal, este esquema tiene orden de convergencia fuerte $\gamma = 0.5$ y orden de convergencia débil $\beta = 1$.

Utilizando otros métodos se puede construir variantes más avanzadas y de órdenes superiores.

Capítulo 5

Modelos estocásticos en finanzas

En esta sección se trabajan los principales modelos basados en EDEs en el campo de las finanzas, desde los más simples hasta los más sofisticados. Además, se desarrolla una pequeña explicación de los conceptos económicos bajo la creación de estos modelos y, después, se validan y visualizan los resultados y soluciones de los modelos a través de la presentación de simulaciones (véase Anexo I) realizadas en el software matemático Mathematica (2023).

5.1 Modelo de Vasicek

El modelo de Vasicek (Vasicek, 1977), desarrollado por Oldrich Vasicek en 1977, fue uno de los primeros modelos de equilibrio para la evolución de los tipos de interés a corto plazo. Captura la idea de la reversión al equilibrio de largo plazo y otros aspectos de la política monetaria, unido a la consideración de los shocks macroeconómicos.

La característica principal de la reversión a la media en el largo plazo es lo que fundamenta que el modelo siga un proceso de Ornstein-Uhlenbeck:

$$dr_t = a(b - r_t)dt + \sigma dW_t, \quad r(0) = r_0$$

donde r_t es el tipo de interés instantáneo (naturaleza continua), $a > 0$ es la velocidad de reversión a la media; $b \geq 0$ es el nivel de equilibrio a largo plazo del tipo de interés; $\sigma > 0$ es

la volatilidad del tipo de interés; y $\{W_t\}_{t \geq 0}$ es un proceso de Wiener.

Observando los coeficientes de deriva $a(b - r_t)$ y de difusión σ , es fácil concluir que ambos verifican las condiciones de Lipschitz y de crecimiento lineal, por lo que se aplica el teorema 3.1.4. Esta EDE tiene solución explícita, como ya se observó en el ejemplo 3.2.9. Para obtenerla, se considera el proceso $Y_t = r_t e^{at}$ y se aplica la fórmula de Itô:

$$dY_t = (ar_t e^{-at} + e^{-at} a(b - r_t))dt + e^{at} \sigma dW_t = abe^{at} dt + e^{at} \sigma dW_t$$

Integrando en $[0, t]$:

$$\int_0^t dY_s = \int_0^t abe^{as} ds + \int_0^t \sigma e^{as} dW_s.$$

Desarrollando las integrales:

$$Y_t - Y_0 = b(e^{at} - 1) + \sigma \int_0^t e^{as} dW_s.$$

Ahora, sustituyendo $Y_t = r_t e^{at}$

$$e^{at} r_t - r_0 = b(e^{at} - 1) + \sigma \int_0^t e^{as} dW_s,$$

y despejando r_t :

$$r_t = e^{-at} \left[r_0 + b(e^{at} - 1) + \sigma \int_0^t e^{as} dW_s \right].$$

Por tanto, la solución explícita para el tipo de interés r_t es:

$$r_t = r_0 e^{-at} + b(1 - e^{-at}) + \sigma \int_0^t e^{-a(t-s)} dW_s.$$

Esta expresión muestra que el proceso r_t sigue una distribución normal con:

$$\mathbb{E}[r_t] = r_0 e^{-at} + b(1 - e^{-at})$$

$$\text{Var}[r_t] = \frac{\sigma^2}{2a}(1 - e^{-2at})$$

De ello se tiene,

$$\lim_{t \rightarrow \infty} \mathbb{E}[r_t] = b$$

$$\lim_{t \rightarrow \infty} \text{Var}[r_t] = \frac{\sigma^2}{2a}$$

Así se demuestra la razón de la reversión a la media, ya que el valor esperado es el nivel de

equilibrio a largo plazo del tipo de interés b . A su vez, se muestra que la varianza a largo plazo es proporcional a la volatilidad del tipo de interés e inversamente proporcional a la velocidad a .

Denotando $B(t, T)$ al valor del bono cupón cero con fecha de vencimiento en T en el tiempo t y considerando el intervalo de tiempo $[t, T]$, se conoce el valor presente o descontado del bono:

$$B(t, T) = B(T, T)e^{-R(t, T)(T-t)}$$

donde $R(t, T)$ es el tipo de interés spot. Esta variable tiene una interpretación económica clara, ya que refleja las expectativas del mercado sobre la evolución futura de los tipos de interés a corto plazo (más una prima de riesgo) y representa el tipo de interés efectivo por unidad de tiempo para una inversión sin riesgo que comienza en t y finaliza en T . Por ello, la función $R(t, T)$ para distintos valores de T , manteniendo t fijo, define una estructura temporal de tipos de interés o curva de rendimientos (*yield curve*), es decir, muestra cómo varían los tipos de interés en función del plazo de vencimiento. Despejando $R(t, T)$ de la anterior expresión, se obtiene la ecuación de dicha curva:

$$R(t, T) = -\frac{\ln B_t}{T-t}$$

Una de las características más importantes del modelo de Vasicek es que permite expresar los precios de los bonos cupón cero. Considerando $B(T, T) = 1$, el precio del bono definido por el modelo de Vasicek es:

$$B(t, T) = A(t, T)e^{-P(t, T)r_t}$$

donde:

$$P(t, T) = \frac{1}{a}(1 - e^{-a(T-t)})$$

$$A(t, T) = \exp \left[\left(b - \frac{\sigma^2}{2a^2} \right) (P(t, T) - (T-t)) - \frac{\sigma^2}{4a} P(t, T)^2 \right]$$

De esta forma, el modelo de Vasicek implica una estructura temporal de tipos de interés:

$$R(t, T) = -\frac{\ln B(t, T)}{T-t} = -\frac{\ln (A(t, T)e^{-P(t, T)r_t})}{T-t} = \frac{-\ln A(t, T)}{T-t} + \frac{P(t, T)}{T-t}r_t.$$

Esto demuestra que, en el modelo de Vasicek, el tipo spot es una función afín del tipo instantáneo r_t . Esta estructura temporal puede generar diversas formas (creciente, decreciente o con joroba) dependiendo de la relación entre el tipo de interés actual r_t y el nivel de largo plazo b . Esta flexibilidad es una de las razones por las que el modelo de Vasicek fue importante en el desarrollo de la modelización de tipos de interés. Aun así, este modelo presenta varias limitaciones:

- Permite tipos de interés negativos. En el momento en el que se publicó este modelo, nunca se había concebido la idea de tipos de interés negativos, hecho que ocurre más tarde tras la crisis global 2008.
- La volatilidad del tipo de interés es constante, independientemente del nivel del tipo de interés.
- El modelo es unifactorial, lo que limita las formas que puede adoptar la estructura temporal.
- La calibración a la estructura temporal observada en el mercado puede ser problemática.

5.2 Modelo de Black-Scholes

Los derivados financieros son instrumentos financieros cuyo valor está determinado por otro activo financiero (acciones, materias primas, criptomonedas, tipos de interés, etc.), que se denomina activo subyacente. Los derivados financieros son contratos sobre pagos o entregas a futuro con ciertas condiciones (Highman J., 2004). Estas condiciones son las que definen el tipo de derivado:

- Contrato forward y contrato a futuro: acuerdo entre dos agentes sobre la venta o compra de un activo en un tiempo futuro determinado por un precio de entrega determinado. Cabe destacar, que los contratos forward, a diferencia de los futuros, no se negocian en el mercado sino que son acuerdos privados entre instituciones o corporaciones financieras (“over the counter market” o mercado extrabursatil)

- Swaps: contratos que regulan un intercambio entre dos agentes de una serie de flujos monetarios en fechas futuras con una cierta periodicidad.
- Opciones: contrato que ofrece a su poseedor (comprador de la opción) el derecho, pero no la obligación, de comprar (call) o vender (put) el activo subyacente a un precio prescrito (strike o precio de ejercicio) dentro de un periodo de tiempo específico. Para obtener este derecho, el poseedor de la opción ha pagado una determinada cantidad de dinero, denominada prima, al escritor (vendedor) de la opción. La fecha de vencimiento es aquella en la que la opción caduca. Existen distintos tipos de opciones, siendo las más conocidas las opciones europeas y opciones americanas. Las opciones europeas solo se pueden ejercer en la fecha de vencimiento, mientras que las opciones americanas pueden ejercerse en cualquier instante hasta la fecha de vencimiento. Por otro lado, las opciones pueden establecerse sobre un activo subyacente o sobre una cesta de activos subyacentes.

Además, dependiendo de la complejidad del contrato, se habla de derivados vanilla (e.g. opciones europeas) o exóticos (e.g. opciones sobre cesta de activos). En esta sección solo se van a tratar las opciones europeas.

La utilidad de opciones financieras está relacionada con la especulación y la cobertura (Vázquez, 2010). La cobertura es una forma de proteger la inversión minimizando el riesgo de pérdidas, es decir, en el uso de opciones, poder tomar posiciones contrarias sobre el activo subyacente. El ejemplo más simple es, ante la posesión de un activo, escribir (vender) una opción call (de compra) sobre el activo, que permite cubrir las pérdidas de la posible devaluación del subyacente, ya que se recibe la prima del comprador.

La especulación con opciones se fundamenta en la posibilidad de apalancamiento, es decir, poder endeudarse para destinar más cantidad de dinero a la inversión y, de esta forma, obtener mayores ganancias (o pérdidas). La idea se basa en conseguir una diferencia a favor entre el strike y el precio de mercado del activo subyacente, de forma que, tras el ejercicio de la opción, se aprovecha inmediatamente la discordancia con el mercado.

Una opción call europea da al comprador el derecho, aunque no la obligación, de comprar el activo subyacente a un precio predeterminado (strike) en una fecha de vencimiento concreta.

El vendedor o escritor de la opción call tiene la obligación de vender el activo si el comprador ejerce su derecho de compra. Entonces, en la fecha de vencimiento puede haber 3 situaciones:

1. El precio de mercado del activo subyacente es menor que el strike (*out-of-the-money*), luego no tiene sentido que el comprador ejerza su derecho de compra.
2. El precio de mercado del activo subyacente es igual al strike (*at-the-money*), luego el comprador puede ejercer su derecho de compra.
3. El precio de mercado del activo subyacente es mayor que el strike (*in-the-money*), luego es lógico que el comprador ejerza su derecho de compra.

En el último caso, la ganancia del comprador ha sido la diferencia entre el precio de mercado del activo subyacente (S) menos el strike (K), mientras que en los otros dos casos ha sido nula (en el caso 2, el comprador podía haber participado directamente en el mercado con el mismo resultado). Por ello, se puede describir la función de pago o payoff de una opción call europea con fecha de vencimiento T como:

$$C(S, T) = C(S_T) = \max(S_T - K, 0) = (S_T - K)^+.$$

Por otro lado, una opción put europea da al poseedor el derecho, pero no la obligación, de venta del activo subyacente a un precio fijo en una fecha de vencimiento determinada. El escritor de la opción put tiene la obligación de comprar el activo subyacente en el caso de que el comprador de la opción decida ejercer su derecho de venta. Al igual que antes, en la fecha de vencimiento puede haber tres situaciones:

1. El precio de mercado del activo subyacente es menor que el strike (*in-the-money*), luego es lógico que el poseedor ejerza su derecho de venta.
2. El precio de mercado del activo subyacente es igual al strike (*at-the-money*), luego el poseedor puede ejercer su derecho de venta.
3. El precio de mercado del activo subyacente es mayor que el strike (*out-of-the-money*), luego no tiene sentido que el poseedor ejerza su derecho de venta.

Esto demuestra que la opción put es la posición contraria a la opción call, luego su ganancia viene determinada por $K - S_T$ en la fecha de vencimiento. Por tanto, la función de pago de una opción put europea en la fecha de vencimiento T se puede expresar como:

$$P(S, T) = P(S_T) = \max(K - S_T, 0) = (K - S_T)^+.$$

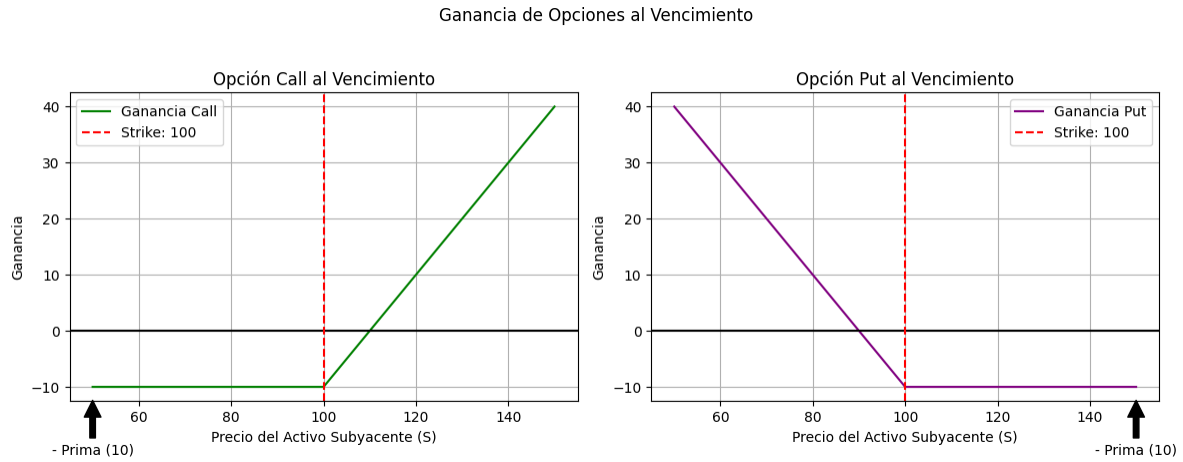


Figura 5.2.1: Opciones europeas a la fecha de vencimiento.

Se observa que las opciones ofrecen una ventaja para el poseedor de esta. Para obtener el derecho de compra o venta de una opción call o put, el comprador debe pagar al escritor de la opción una prima, que determina el valor de la opción. Es lógico asumir que el valor de la opción está definido por su función de pago, que en el caso de las opciones europeas se ha determinado que es:

$$V(S, T) = \begin{cases} (S_T - K)^+ & \text{(Opción call)} \\ (K - S_T)^+ & \text{(Opción put)} \end{cases}$$

Se sabe que la función $V(S, T)$ es valor (precio) de la opción en el instante futuro T . Pero, para la compraventa de opciones, la cuestión importante es conocer cuál es el precio justo $V(S, t)$ de una opción para $t < T$. Esta cuestión es no-trivial debido a que el valor de un activo de riesgo es aleatorio, luego el valor S no es conocido cuando se realiza la compraventa de la opción en $t = 0$.

En 1973, se publica el modelo de Black-Scholes (Black y Scholes, 1973), desarrollado por

Fischer Black y Myron Scholes, y Robert Merton de forma paralela, que constituye uno de los avances más significativos en la matemática financiera. Este modelo proporciona un marco teórico para la valoración de opciones y otros derivados financieros.

Antes de formular el modelo de Black-Scholes, se deben de aclarar las bases en las que se fundamenta. Lo primero, se distingue entre activos de riesgo (acciones, materias primas, etc.) y activos sin riesgo, considerando solo los bonos. Si se invierte $B(t_0)$ en el instante t_0 en un bono con un tipo de interés libre de riesgo $r > 0$, entonces el valor del bono en un instante $t \geq t_0$ es:

$$B(t) = e^{r(t-t_0)} B(t_0);$$

es decir, se asume un pago de interés continuo.

Relacionado con esta idea, aparece el concepto de arbitraje (Jahnke, 2019). Esto se refiere a la existencia de un portfolio, un conjunto de activos financieros o inversiones, que no requieren una inversión inicial y aseguran retornos no negativos. En el caso de los bonos, siempre ofrecen un retorno igual a r , pero requieren de una inversión inicial. La ausencia de arbitraje significa que no es posible la obtención de una ganancia instantánea libre de riesgo, o que esta oportunidad no podría mantenerse durante un período significativo de tiempo.

Por otro lado, se asume que el valor de los activos de riesgo S varía de forma aleatoria según las hipótesis de mercado eficiente:

- El valor actual del activo es consecuencia de su valor en el pasado inmediato.
- El mercado del activo responde de manera instantánea a cualquier nueva información sobre el mismo.

De esta forma, suponiendo que el valor del activo en el instante t como S_t , la evolución de su precio puede expresarse mediante la siguiente ecuación diferencial estocástica:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

donde $\mu \in \mathbb{R}$ es la tendencia, $\sigma \in \mathbb{R}$ es la volatilidad y $\{W_t\}_{t \geq 0}$ es un proceso de Wiener escalar. Es decir, el precio del activo se modela con un movimiento browniano geométrico, EDE lineal estudiada anteriormente. Como se mostró en el ejemplo 3.2.8, la solución explícita

de esta ecuación es:

$$S_t = S_0 \exp [(\mu - \sigma^2)/2 + \sigma X_t].$$

Por la definición del movimiento browniano geométrico, se establece que S_t es log-normal:

$$\ln(S_t) = \ln(S_{t_0}) + at + \sigma W_t \sim \mathcal{N}(\ln(S_{t_0}) + at, \sigma^2 t).$$

Una vez definidos estos conceptos, las suposiciones generales que se van a tomar son (Jahnke, 2019):

- H1. El precio del activo subyacente S sigue un movimiento Browniano geométrico. Se supone $\sigma, \mu \in \mathbb{R}$ constante y conocidas en el tiempo.
- H2. Existe un tipo de interés libre de riesgo $r > 0$ y se asumen pagos de interés continuos.
- H3. Principio de ausencia arbitraje. Todos los portfolios libres de riesgo tienen mismo retorno igual a r .
- H4. Mercado líquido: no existen costes de transacción, impuestos, etc. Se puede comprar y vender activos o fracciones de estos en cualquier momento (tiempo continuo). La venta de un activo no afecta significativamente a su valor.
- H5. Un agente puede vender activos que no posea (*short selling*).
- H6. El activo subyacente no paga dividendos.

A partir de estas hipótesis, se deriva la esencia del modelo de Black-Scholes. Se considera el portfolio Π (autofinanciado) para a_t unidades del activo subyacente S_t y b_t unidades del bono B_t :

$$\Pi_t = s(t)S_t + b(t)B_t.$$

Por otro lado, se considera una opción $V_T = V(S, T)$, de forma que por su valor en T se tiene:

$$a_T S_T + b_T B_T = V_T.$$

Por el principio de no arbitraje: $a_t S_t + b_t B_t = V_t$ para cada $0 \leq t \leq T$. Entonces,

$$dV_t = a_t dS_t + b_t dB_t.$$

Considerando $V_t = V(S_t, t)$, $t \in [0, T]$, se aplica la fórmula de Itô:

$$dV_t = \left(\frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} \right) dt + \sigma S_t \frac{\partial V}{\partial S_t} dW_t$$

Como $dS_t = \mu S_t dt + \sigma S_t dW_t$ y $dB_t = r B_t dt$, se tiene que:

$$dV_t = a_t(\mu S_t dt + \sigma S_t dW_t) + b_t(r B_t dt) = (a_t \mu S_t + b_t r B_t) dt + a_t \sigma S_t dW_t$$

Igualando los coeficientes de deriva y difusión:

$$\begin{aligned} a_t \sigma S_t &= \sigma S_t \frac{\partial V}{\partial S_t} \\ a_t \mu S_t + b_t r B_t &= \frac{\partial V}{\partial t} + \mu S_t \frac{\partial V}{\partial S_t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2}, \end{aligned}$$

y despejando

$$\begin{aligned} a_t &= \frac{\partial V}{\partial S_t}, \quad S_t > 0 \\ b_t &= \frac{\left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} \right)}{r B_t}, \quad r B_t \neq 0. \end{aligned}$$

Sustituyendo

$$V_t = a_t S_t + b_t B_t = \frac{\partial V}{\partial S_t} S_t + \frac{\left(\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} \right)}{r B_t} B_t.$$

Y reordenando, se obtiene la conocida ecuación de Black-Scholes:

$$\frac{\partial V_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} + r S_t \frac{\partial V_t}{\partial S_t} - r V_t = 0$$

Con esto, se ha construido una ecuación en derivadas parciales determinista para la valoración de una opción. Para una opción de compra europea (call) con precio de ejercicio K y vencimiento T , la solución analítica es:

$$C(S, t) = S \mathcal{N}(d_1) - K e^{-r(T-t)} \mathcal{N}(d_2),$$

y para una opción de venta europea (put):

$$P(S, t) = K e^{-r(T-t)} \mathcal{N}(-d_2) - S \mathcal{N}(-d_1),$$

donde:

$$d_1 = \frac{\ln(S/K) + (r + \sigma^2/2)(T - t)}{\sigma\sqrt{T - t}}$$

$$d_2 = d_1 - \sigma\sqrt{T - t},$$

y $\mathcal{N}(x)$ es la función de distribución acumulada de una variable normal estándar. El modelo de Black-Scholes descompone el valor de una opción en dos componentes:

- $S\mathcal{N}(d_1)$: representa el valor esperado de recibir el activo si la opción termina en dinero.
- $Ke^{-r(T-t)}\mathcal{N}(d_2)$: representa el valor presente del precio de ejercicio ponderado por la probabilidad de ejercicio.

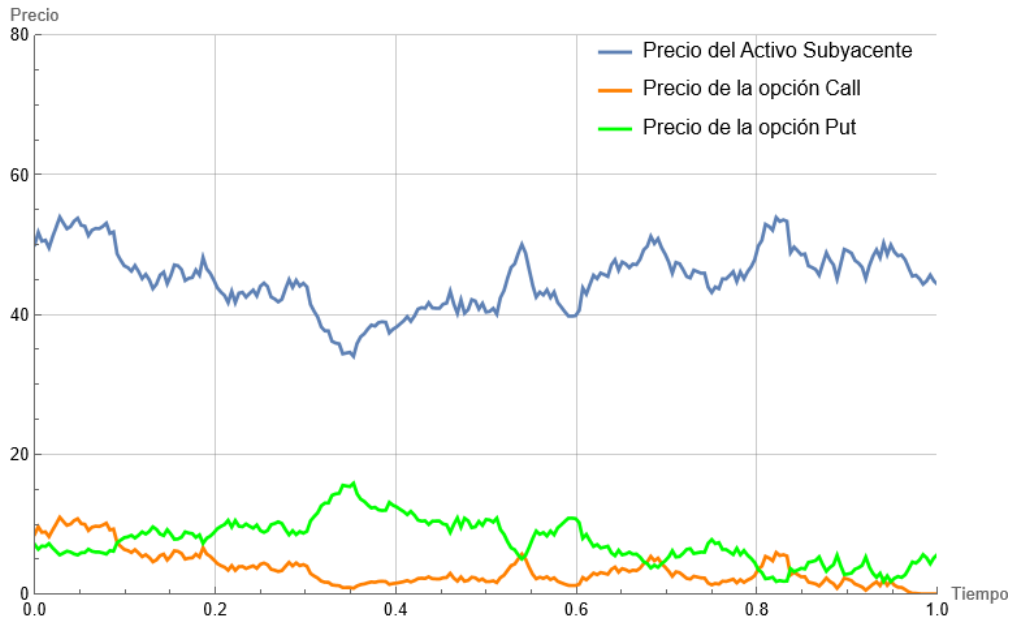


Figura 5.2.2: Trayectoria del precio del activo subyacente, opción call y opción put.

En la figura 5.2.2, se muestra una simulación de la evolución del valor de una opción call y put y la trayectoria del precio subyacente con los parámetros

$$S_0 = K = 50, \quad \mu = 0.02, \quad \sigma = 0.4, \quad r = 0.03, \quad T = 1, \quad h = 1/252.$$

Se observa claramente la dinámica contraria entre la opción call y la opción put: mientras que la opción call sigue, en cierta medida, la dirección del precio del activo, la opción put se mueve en dirección contraria. Además, se resalta la idea de que, al acercarse a la fecha de

vencimiento $T = 1$, el valor de las opciones tiende hacia su valor intrínseco. En el vencimiento, $S_T < K$, por ello, la opción call vale 0, mientras que la opción put $K - S_T$.

La relación de las tres trayectorias satisface la paridad put-call:

$$S_t - C_t + P_t = Ke^{-r(T-t)}$$

donde C_t es el valor de la opción call y P_t el valor de la opción put.

A pesar de su importancia, el modelo de Black-Scholes presenta varias desventajas:

- Volatilidad constante: En la práctica, se observa que la volatilidad varía con el tiempo y con el nivel de precios del activo (sonrisa de volatilidad).
- Distribución log-normal: Los rendimientos reales de los activos exhiben colas más pesadas (asimetrías) que las predichas por una distribución normal.
- Saltos en los precios: El modelo no puede capturar movimientos bruscos o discontinuidades en los precios de los activos.
- Supuestos de liquidez y costos de transacción: Los mercados reales presentan fricciones que el modelo ignora.

Estas limitaciones han motivado el desarrollo de modelos más sofisticados, como el modelo de Merton de salto-difusión y los modelos de volatilidad estocástica.

5.3 Modelo de salto-difusión

Robert Merton extendió el modelo de Black-Scholes para incorporar saltos discretos en los precios de los activos (Merton, 1976), reconociendo que los mercados financieros frecuentemente experimentan movimientos bruscos debido a la llegada repentina de nueva información.

En el modelo de Merton de salto-difusión, la dinámica del precio del activo se describe mediante:

$$dS_t = (\mu - \lambda\kappa)S_t dt + \sigma S_t dW_t + (Y_t - 1)S_t dN_t$$

donde $\mu \in \mathbb{R}$ y $\sigma \in \mathbb{R}$ son los parámetros de tendencia y volatilidad; W_t es un proceso de Wiener escalar; N_t es un proceso de Poisson escalar con intensidad $\lambda > 0$, que representa la frecuencia de los saltos; $\kappa = \mathbb{E}[Y_t - 1]$ con Y_t variables aleatorias independientes e idénticamente distribuidas que representan el tamaño de los saltos (normalmente se asumen log-normal).

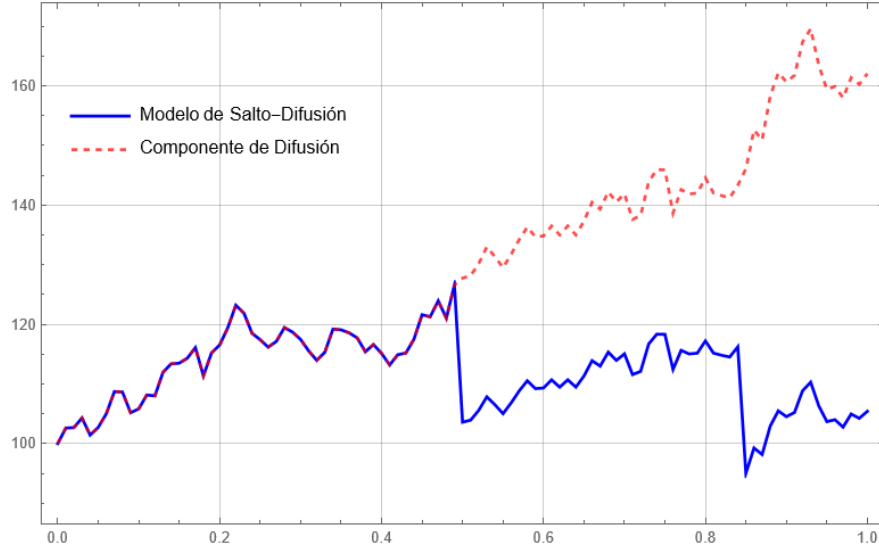


Figura 5.3.1: Trayectoria del modelo de salto-difusión de Merton frente a su término difusivo.

El modelo de Merton incluye un proceso de Poisson N_t que introduce discontinuidades, como se observa en la figura 5.3.1. Esta característica viola fundamentalmente las condiciones de Lipschitz global y crecimiento lineal porque los saltos crean trayectorias discontinuas y estas condiciones están diseñadas para EDEs con coeficientes continuos. No obstante, existen teoremas de existencia y unicidad para EDE con saltos que garantizan la unicidad de la solución bajo condiciones adecuadas (Przybyłowicz y Szölgyenyi, 2021).

Por tanto, el modelo de Merton permite capturar eventos extremos y modelar el riesgo de eventos raros pero significativos (como crisis financieras). Además, verifica las colas pesadas en la distribución de rendimientos (Gugole, 2016). Por contra, la calibración del modelo es más compleja debido al mayor número de parámetros y la volatilidad de la componente de difusión sigue siendo constante, sin capturar adecuadamente la dinámica temporal de esta.

El precio de una opción bajo el modelo de Merton satisface:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV + \lambda \mathbb{E}[V(Y_t S_t, t) - V(S_t, t)] - \lambda \kappa S_t \frac{\partial V}{\partial S_t} = 0$$

La fórmula de valoración para una opción europea es:

$$V(S, t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda'(T-t)} [\lambda'(T-t)]^n}{n!} V_{BS}(S, t; \sigma_n)$$

donde

$$\lambda' = \lambda(1 + \kappa), \quad \sigma_n^2 = \sigma^2 + \frac{n\delta^2}{T-t}$$

y $V_{BS}(S, t; \sigma_n)$ es el precio de la opción según el modelo de Black-Scholes.

5.4 Modelo CIR

El modelo de Cox-Ingersoll-Ross (CIR), creado por John Cox, Jonathan Ingersoll y Stephen Ross en 1985, aborda algunas de las limitaciones del modelo de Vasicek, especialmente la posibilidad de tipos de interés negativos. El modelo CIR (Cox et al., 2005) sigue una estructura parecida al modelo de Vasicek. El tipo de interés instantáneo r_t sigue el proceso:

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t, \quad r_0 \geq 0$$

donde $\{W_t\}_{t \geq 0}$ es un proceso de Wiener escalar y los parámetros a, b y σ son constantes positivas y tienen interpretaciones similares a las del modelo de Vasicek. La diferencia fundamental está en el término de difusión, que ahora depende de $\sqrt{r_t}$.

El modelo CIR presenta varias propiedades deseables, entre ellas la no negatividad de los tipos de interés (c.p.t). Incluso, se demuestra la positividad estricta, fundamentada en la condición de Feller:

Proposición 5.4.1. *Si a, b, σ^2 son positivos y satisfacen la condición $2ab > \sigma^2$ entonces la solución r_t del modelo CIR con $r_0 > 0$ verifica que $r_t > 0$, $\forall t \in \mathbb{R}^+$, c.p.t.*

Al igual que el modelo de Vasicek, conserva la propiedad de reversión a la media, hacia el nivel de equilibrio b a largo plazo. En cambio, este modelo añade el factor de la volatilidad σ dependiente del nivel r_t , lo que captura la idea de que los tipos de interés más altos tienden a ser más volátiles.

Debido a la introducción de este término $\sqrt{r_t}$ en el coeficiente de difusión, este no satisface la condición de Lipschitz y, por tanto, no se puede aplicar el teorema 3.1.4. No obstante, existen una base teórica refinada y resultados específicos que demuestran la existencia y unicidad de la solución. La solución explícita del modelo CIR no está disponible de forma general, pero se han obtenido resultados cumpliéndose cierta relación entre los parámetros de la ecuación.

Proposición 5.4.2. *Dada la ecuación*

$$dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t, \quad r_0 = c,$$

si los parámetros verifican

$$4ab = \sigma^2,$$

entonces, la única solución exacta viene dada por

$$r_t = e^{-at} \left(\sqrt{c} + \sqrt{ab} \int_0^t e^{\frac{as}{2}} dW_s \right)^2.$$

La solución anterior se escribe $r_t = Z_t^2$ donde

$$Z_t = e^{\frac{-at}{2}} \sqrt{c} + e^{\frac{-at}{2}} \sqrt{ab} \int_0^t e^{\frac{as}{2}} dW_s$$

es una distribución normal de media

$$\mu(t) = e^{\frac{-at}{2}} \sqrt{c}$$

y varianza

$$\begin{aligned} s^2(t) &= \mathbb{E} \left[\left(e^{\frac{-at}{2}} \sqrt{ab} \int_0^t e^{\frac{as}{2}} dW_s \right)^2 \right] = e^{-at} ab \mathbb{E} \left[\left(\int_0^t e^{\frac{as}{2}} dW_s \right)^2 \right] \\ &= e^{-at} ab \int_0^t e^{as} ds = b(1 - e^{-at}). \end{aligned}$$

Luego $\frac{1}{s(t)}Z_t \sim \mathcal{N}\left(\frac{\mu(t)}{s(t)}, 1\right)$ y $r_t = Z_t^2$ se distribuye como una $s^2(t)$ multiplicada por una variable chi-cuadrado no central con 1 grado de libertad y parámetro de no centralidad

$$\lambda(t) = \left(\frac{\mu(t)}{s(t)} \right)^2 = \frac{c}{b(e^{at} - 1)},$$

es decir,

$$\mathbb{P}(r_t \leq x | r_0 = 0) = F_{\chi^2} \left(\frac{x}{b(1 - e^{-at})}; 1, \lambda(t) \right).$$

Tomando la partición discreta del intervalo temporal con un paso constante h y la solución exacta presentada en la proposición 5.4.2, se tiene

$$X_{n+1} = e^{-ah} \left(\sqrt{X_n} + \sqrt{ab} \Delta V_n \right)^2$$

donde

$$\Delta V_n = \int_{t_n}^{t_{n+1}} e^{\frac{a(s-t_n)}{2}} dW_s \sim \mathcal{N} \left(0, \sqrt{(e^{ah} - 1)/a} \right).$$

Luego, dadas las propiedades del proceso de Wiener estándar, se expresa la solución exacta como

$$X_{n+1} = e^{-ah} \left(\sqrt{X_n} + \sqrt{\frac{b(e^{ah} - 1)}{h}} \Delta W_n \right)^2$$

Dado que solo se conoce la solución exacta cuando se satisface la relación $4ab = \sigma^2$, es necesario recurrir a esquemas numéricos para integrarla en el resto de los casos. Los métodos numéricos que utilizan evaluaciones sobre el coeficiente de difusión $\sigma\sqrt{r_t}$ o sus derivadas, no están bien definidos si pueden aparecen valores negativos. Por esto, se han propuesto distintos esquemas especialmente diseñados para solventar esta cuestión y mantener las propiedades cualitativas de la solución exacta. En concreto, se van a presentar métodos que aproximan la solución preservando la propiedad de no negatividad de la ecuación del modelo y la reversión a la media.

En (Higham y Mao, 2005), los autores estudian la modificación del método de Euler (HM) dada por

$$Y_{n+1} = Y_n + a(b - Y_n)h + \sigma\sqrt{|Y_n|}\Delta W_n$$

demostrando que converge en el sentido fuerte y débil al modelo CIR. En concreto, los autores también estudiaron la preservación de la reversión a la media, concluyendo que verifica dicha propiedad cuando $h < 2/a$. En (Deelstra y Delbaen, 1998), los autores proponen el método de Euler Truncado (TE)

$$Y_{n+1} = Y_n + a(b - Y_n)h + \sigma\sqrt{\max\{Y_n, 0\}}\Delta W_n.$$

Alfonsi desarrolla varios métodos numéricos especialmente diseñados para resolver el modelo CIR (Alfonsi, 2005). Cuando

$$\sigma^2 \leq 4ab$$

definió el método Drift Implícito (DI)

$$Y_{n+1} = \left(\frac{\frac{\sigma}{2} \Delta W_n + \sqrt{Y_n} + \sqrt{\left(\frac{\sigma}{2} \Delta W_n + \sqrt{Y_n}\right)^2 + 4 \left(1 + a \frac{h}{2}\right) \frac{ab - \frac{\sigma^2}{4} h}}}{2 \left(1 + a \frac{h}{2}\right)} \right)^2.$$

Se mostró que este esquema tiene orden fuerte $\gamma = 1/2$ bajo la hipótesis $\sigma^2 < 2ab$ (Dereich et al., 2012). Posteriormente, Alfonsi probó que su orden fuerte es 1 bajo la condición más restrictiva $\sigma^2 < ab$ (Alfonsi, 2013). El mismo autor propuso en (Alfonsi, 2005) el esquema llamado $E(0)$

$$Y_{n+1} = \left(\left(1 - \frac{ah}{2}\right) \sqrt{Y_n} + \frac{\sigma \Delta W_n}{2 \left(1 - \frac{ah}{2}\right)} \right)^2 + \left(ab - \frac{\sigma^2}{4}\right) h,$$

definido cuando se cumple $\sigma^2 \leq 4ab$ y $ah \neq 2$. Se demostró que tiene un orden de convergencia fuerte logarítmico y alcanza un orden de convergencia débil de 1.

Otro esquema interesante es el método θ -Milstein, el cual aplicado al modelo CIR resulta

$$Y_{n+1} = (1 - ah + a\theta h) Y_n - a\theta h Y_{n+1} + abh + \sigma \sqrt{Y_n} \Delta W_n + \frac{\sigma^2}{4} (\Delta W_n^2 - h),$$

que se puede expresar de forma explícita como

$$Y_{n+1} = \frac{1}{1 + a\theta h} \left\{ (1 - ah + a\theta h) Y_n + \left(ab - \frac{\sigma^2}{4}\right) h + \sigma \sqrt{Y_n} \Delta W_n + \frac{\sigma^2}{4} \Delta W_n^2 \right\}.$$

Se puede demostrar que el esquema θ -Milstein preserva la no negatividad de la solución exacta (Llamazares-Elias y Tocino, 2024).

Proposición 5.4.3. *El esquema θ -Milstein*

$$Y_{n+1} = \frac{1}{1 + a\theta h} \left\{ (1 - ah + a\theta h) Y_n + \left(ab - \frac{\sigma^2}{4}\right) h + \sigma \sqrt{Y_n} \Delta W_n + \frac{\sigma^2}{4} \Delta W_n^2 \right\},$$

con $\theta \geq 1$ y $Y_0 \geq 0$ preserva la no-negatividad de la solución exacta r_t si los parámetros del

modelo CIR verifican la condición

$$\sigma^2 \leq 4ab$$

En concreto, si los parámetros verifican la condición de Feller, entonces preserva la positividad de la solución exacta.

Demostración. El esquema θ -Milstein expresado de forma explícita se puede escribir como

$$Y_{n+1} = \frac{1}{1 + a\theta h} \left\{ ah(\theta - 1)Y_n + \left(ab - \frac{\sigma^2}{4} \right) h + g(\Delta W_n) \right\},$$

donde $g(z) = Y_n + \sigma\sqrt{Y_n}z + \frac{\sigma^2}{4}z^2$ para $z \in \mathbb{R}$. La función cuadrática g alcanza su mínimo absoluto en $z^* = -2\sqrt{Y_n}/\sigma$. Entonces, $\forall z \in \mathbb{R}$

$$g(z) \geq g(z^*) = Y_n + \sigma\sqrt{Y_n} \frac{-2\sqrt{Y_n}}{\sigma} + \frac{\sigma^2}{4} \left(\frac{-2\sqrt{Y_n}}{\sigma} \right)^2 = 0.$$

Luego, usando este resultado en la expresión inicial, se tiene

$$Y_{n+1} \geq \frac{1}{1 + a\theta h} \left\{ ah(\theta - 1)Y_n + \left(ab - \frac{\sigma^2}{4} \right) h \right\} \geq 0$$

si se cumple que $\theta \geq 1$, $Y_0 \geq 0$ y $\sigma^2 \leq 4ab$. Fácilmente se observa que, bajo la condición de Feller, $Y_{n+1} > 0$. □

Por otro lado, se puede probar que, bajo las hipótesis de la proposición 5.4.3, los esquemas θ -Milstein tienen un orden de convergencia fuerte logarítmico y orden de convergencia débil de 1. Además, los esquemas θ -Milstein capturan la propiedad de reversión a la media (Llamazares-Elias y Tocino, 2024).

Teorema 5.4.4. *Dado cualquier $h > 0$, el esquema θ -Milstein con $\theta \geq 1$ preserva la media a largo plazo de la solución exacta de cualquier modelo CIR con $4ab - \sigma^2 \geq 0$.*

Demostración. Considérese el esquema θ -Milstein

$$Y_{n+1} = \frac{1}{1 + a\theta h} \left\{ (1 - ah + a\theta h) Y_n + \left(ab - \frac{\sigma^2}{4} \right) h + \sigma\sqrt{Y_n}\Delta W_n + \frac{\sigma^2}{4}\Delta W_n^2 \right\}.$$

Sabiendo que $\mathbb{E}[\Delta W_n] = 0$ y $\mathbb{E}[\Delta W_n^2] = h$, tomando esperanzas

$$\mathbb{E}[Y_{n+1}] = A\mathbb{E}[Y_n] + B,$$

donde

$$A = \frac{1 - ah(1 - \theta)}{1 + a\theta h}, \quad B = \frac{abh}{1 + a\theta h}.$$

Entonces,

$$\mathbb{E}[Y_n] = A^n \mathbb{E}[Y_0] + \frac{1 - A^n}{1 - A} B = A^n (\mathbb{E}[Y_0] - \mu) + \mu,$$

donde se ha usado que $B/(1 - A) = \mu$. Como $\theta \geq 1$,

$$0 < \frac{1 + \alpha\Delta(\theta - 1)}{1 + \alpha\theta\Delta} = A = 1 - \frac{\alpha\Delta}{1 + \alpha\theta\Delta} < 1.$$

Luego $A^n \rightarrow 0$ cuando $n \rightarrow +\infty$, que implica

$$\lim_{n \rightarrow +\infty} \mathbb{E}[Y_n] = b.$$

□

Para estudiar la propiedad de la media a largo plazo (Llamazares-Elias y Tocino, 2024), se ha simulado $N = 10^5$ trayectorias de las soluciones numéricas dadas por los esquemas θ -Milstein ($\theta = 1$), HM, TE, DI y E(0) sobre el intervalo $[0, 15]$ con $h = 1/8$ los parámetros

$$a = 0.5, \quad b = 0.5, \quad \sigma = 1, \quad r_0 = 0.21,$$

Nótese que $\sigma^2 = 4ab$, $h < 2/a$. Para cada método se ha calculado cada punto medio $\bar{Y}_n$ y la distancia $d_Y(t_n)$ entre $\bar{Y}_n$ y la media a largo plazo como

$$\bar{Y}_n = \frac{1}{N} \sum_{i=1}^N Y_{n,i}; \quad d_Y(t_n) = |\bar{Y}_n - b|$$

En la figura 5.4.1 se presentan los resultados sobre la trayectoria media y la evolución de la distancia respecto a la media a largo plazo para cada esquema. Se observa que el método θ -Milstein y Higham-Mao (Euler modificado) reducen la distancia d_Y progresivamente, luego revierten a la media de largo plazo, de acuerdo con los resultados teóricos. A su vez, parece que lo mismo ocurre con el método de Euler truncado, aunque dicha proposición no está demostrada.

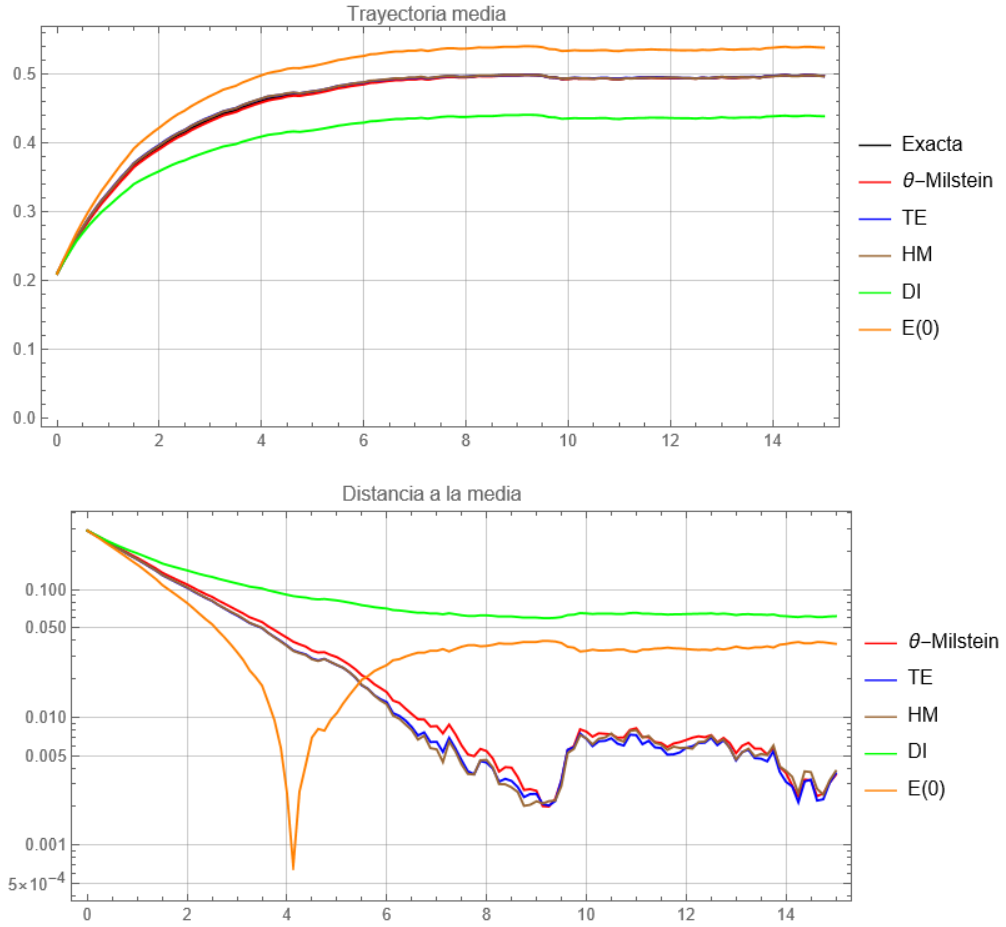


Figura 5.4.1: Trayectoria media y su distancia respecto a la media a largo plazo.

El modelo CIR también define una valoración de los bonos cupón cero:

$$B(t, T) = A(t, T)e^{-P(t, T)r_t}$$

donde:

$$P(t, T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma}$$

$$A(t, T) = \left[\frac{2\gamma e^{(\gamma+a)(T-t)/2}}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma} \right]^{2ab/\sigma^2}$$

$$\gamma = \sqrt{a^2 + 2\sigma^2}$$

De forma que

$$R(t, T) = -\frac{\ln B(t, T)}{T - t} = -\frac{\ln (A(t, T)e^{-P(t, T)r_t})}{T - t} = \frac{-\ln A(t, T)}{T - t} + \frac{P(t, T)}{T - t}r_t$$

al igual que en el modelo Vasicek, el tipo spot es una función afín del tipo instantáneo r_t . Aun así, sigue siendo un modelo unifactorial, lo que limita su capacidad para ajustarse a estructuras temporales complejas.

5.5 Modelos de volatilidad estocástica

Uno de las suposiciones fundamentales de los modelos vistos, es la consideración de una volatilidad constante al describir el movimiento del precio del activo. En la práctica, se observa que la volatilidad de los activos financieros no es constante, sino que fluctúa a lo largo del tiempo y muestra características como la reversión a la media.

El modelo de Heston, propuesto por Steven Heston, considera tanto el precio del activo como su varianza instantánea siguen procesos estocásticos (Heston, 1993). Se establece el sistema:

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dW_t^1 \\ dv_t &= \kappa(\theta - v_t)dt + \sigma_v \sqrt{v_t} dW_t^2 \end{aligned}$$

donde $\mu \in \mathbb{R}$ es el rendimiento esperado del activo; $\kappa \in \mathbb{R}$ es la velocidad de reversión a la media de v_t ; $\theta > 0$ es el nivel de largo plazo de v_t ; σ_v es la volatilidad de la volatilidad; W_t^1 y W_t^2 son movimientos brownianos con correlación ρ : $dW_t^1 dW_t^2 = \rho dt$.

El modelo define la volatilidad como un proceso de reversión a la media, igual que la ecuación del modelo CIR. La correlación entre los dos procesos brownianos permite capturar el “efecto apalancamiento”, esto es, la tendencia de la volatilidad a aumentar cuando los precios del activo caen. En la figura 5.5.1, se visualiza esta propiedad que define la importancia de este modelo.

Al igual que ocurre el modelo CIR, el modelo de Heston no satisface la condición de Lipschitz global ni la de crecimiento lineal debido al término $\sqrt{v(t)}$. Al tratarse de un sistema desacoplado, se puede tratar la ecuación de la volatilidad como en el modelo CIR.

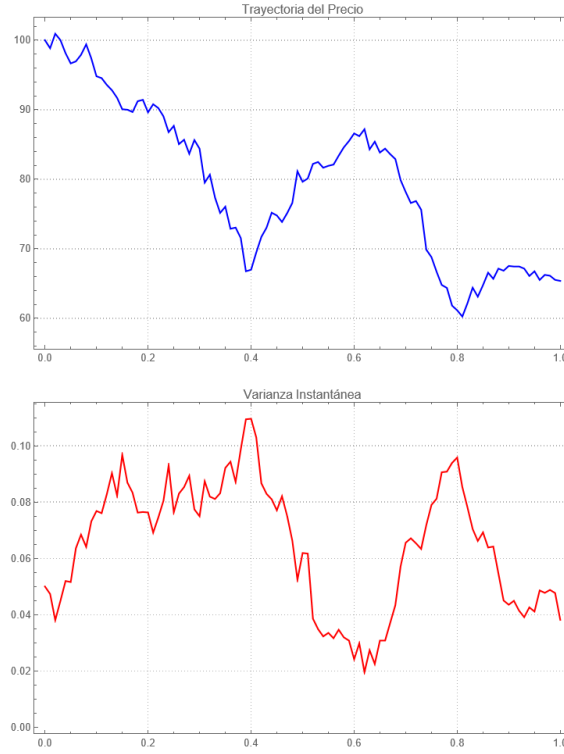


Figura 5.5.1: Trayectoria del modelo de Heston.

El precio de una opción $V(S, v, t)$ bajo el modelo de Heston satisface:

$$\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \rho\sigma_v vS\frac{\partial^2 V}{\partial S\partial v} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 V}{\partial v^2} + rS\frac{\partial V}{\partial S} + \kappa(\theta - v)\frac{\partial V}{\partial v} - rV = 0$$

La fórmula para el precio de una opción de compra europea es:

$$V(S, v, t) = SP_1 - Ke^{-r(T-t)}P_2,$$

donde P_1 y P_2 son probabilidades ajustadas al riesgo que se obtienen mediante la inversión de funciones características:

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[\frac{e^{-i\phi \ln(K)} f_j(S, v, t; \phi)}{i\phi} \right] d\phi, \quad j = 1, 2.$$

Las funciones características f_j tienen expresiones complejas que involucran funciones exponenciales y logarítmicas de los parámetros del modelo.

Bibliografía

- Alfonsi, A. (2005). On the discretization schemes for the cir (and bessel squared) processes. *Monte Carlo Methods Appl.*, 11(4):355–384.
- Alfonsi, A. (2013). Strong order one convergence of a drift implicit euler scheme: Application to the cir process. *Statistics & Probability Letters*, 83(2):602–607.
- Arnold, L. (1974). *Stochastic Differential Equations: Theory and Applications*. John Wiley & Sons.
- Black, F. y Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54.
- Braumann, C. A. (2019). *Introduction to Stochastic Differential Equations with Applications to Modelling in Biology and Finance*. Wiley.
- Cox, J., Ingersoll, J., y Ross, S. (2005). A theory of the term structure of interest rates. En *Theory of valuation*, pág. 129–164. World Scientific.
- Deelstra, G. y Delbaen, F. (1998). Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied Stochastic Models and Data Analysis*, 14(1):77–84.
- Dereich, S., Neuenkirch, A., y Szpruch, L. (2012). An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proceedings of the royal society A: mathematical, physical and engineering sciences*, 468(2140):1105–1115.
- Dunn, R., Hauser, P., Seibold, T., y Gong, H. Estimating option prices with heston’s stochastic volatility model.
- Gugole, N. (2016). Merton jump-diffusion model versus the black and scholes approach for the log-returns and volatility smiling fitting. *International Journal of Pure and Applied Mathematics*, 109(3):719–736.
- Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343.

- Higham, D. J. y Mao, X. (2005). Convergence of Monte Carlo simulations involving the mean-reverting square root process. *Journal of Computational Finance*, 8(3):35–61.
- Highman J., D. (2004). *An Introduction to Financial Option Valuation*. Cambridge University Press, Cambridge, Reino Unido.
- Itô, K. (1944). Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519–524.
- Jahnke, T. (2019). *Numerical methods in mathematical finance*. Karlsruher Institut für Technologie.
- Kloeden, P. y Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations*. Springer.
- Llamazares Elias, S. y Tocino, A. (2022). On the numerical integration of the CIR model. *Available at SSRN 4010731*.
- Llamazares-Elias, S. y Tocino, A. (2024). Preservation of structural properties of the CIR model by θ -Milstein schemes. *arXiv preprint arXiv:2412.17983*.
- Mathematica, W., Versión 13.3.1.0, [Software] (2023). Recuperado de <https://www.wolfram.com/mathematica/>.
- Matsuda, K. (2004). Introduction to merton jump diffusion model.
- Merton, R. (1976). Option pricing when underlying returns are discontinuous. *Journal of Financial Economics*, 3:125–144.
- Przybyłowicz, P. y Szölgyenyi, M. (2021). Existence, uniqueness, and approximation of solutions of jump-diffusion sdes with discontinuous drift. *Applied Mathematics and Computation*, 403(15):126–149.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188.
- Vázquez, C. (2010). An introduction to black-scholes modeling and numerical methods in derivatives pricing. *MAT - Serie A*, (17):1–44.

Anexo I: Código Mathematica

Modelo Black-Scholes

```
(*Parametros*)
S0 = 50;
\[Mu] = 0.02;
\[Sigma] = 0.4;
r = 0.03;
K = 50;
T = 1;
pasos = 252;
dt = T/pasos;

(*Inicio*)
price = Table[0, {pasos + 1}];
price[[1]] = S0;

(*Simulacion precio del activo*)
Do[
  W = RandomVariate[NormalDistribution[0, 1]];
  price[[i + 1]] =
    price[[i]]*Exp[(\[Mu] - 0.5*\[Sigma]^2)*dt + \[Sigma]*Sqrt[dt]*W
    ];
  , {i, 1, pasos}];

(*Simulacion precio de la opcion call y put*)
Nd1[s_, k_, r_, sigma_,
  tau_] := (Log[s/k] + (r + 0.5*sigma^2)*tau)/(sigma*Sqrt[tau]);

Nd2[s_, k_, r_, sigma_, tau_] :=
  Nd1[s, k, r, sigma, tau] - sigma*Sqrt[tau];
```

```
CallPrice[s_, k_, r_, sigma_, tau_] := If[tau <= 0, Max[0, s - k],
  s*CDF[NormalDistribution[0, 1], Nd1[s, k, r, sigma, tau]] -
  k*Exp[-r*tau]*
  CDF[NormalDistribution[0, 1], Nd2[s, k, r, sigma, tau]]];

PutPrice[s_, k_, r_, sigma_, tau_] := If[tau <= 0, Max[0, k - s],
  k*Exp[-r*tau]*
  CDF[NormalDistribution[0, 1], -Nd2[s, k, r, sigma, tau]] -
  s*CDF[NormalDistribution[0, 1], -Nd1[s, k, r, sigma, tau]]];

optionsCall = Table[0, {pasos + 1}];
optionsPut = Table[0, {pasos + 1}];
Do[
  tau = T - i*dt;
  optionsCall[[i + 1]] =
    CallPrice[price[[i + 1]], K, r, \[Sigma], tau];
  optionsPut[[i + 1]] =
    PutPrice[price[[i + 1]], K, r, \[Sigma], tau], {i, 0, pasos}];

timePoints = Table[i*dt, {i, 0, pasos}];
pricedata = Transpose[{timePoints, price}];
optionsCallData = Transpose[{timePoints, optionsCall}];
optionsPutData = Transpose[{timePoints, optionsPut}];
(*Grafico*)
dataplot =
  ListLinePlot[{pricedata, optionsCallData, optionsPutData},
    PlotRange -> {{0.0, 1.0}, {0, 80}},
    AxesLabel -> {"Tiempo", "Precio"}, ImageSize -> Large,
    GridLines -> Automatic,
    PlotLegends ->
      Placed[{"Precio del Activo Subyacente",
```

```
"Precio_de_la_opcion_Call", "Precio_de_la_opcion_Put"}, {0.8,
0.9}], PlotStyle -> {Automatic, Orange, Green} ];
```

Modelo de salto-difusión

```
(*Parametros*)
\[Mu] = 0.1;
\[Sigma] = 0.2;
\[Lambda] = 2;
\[Mu]J = -0.1;
\[Sigma]J = 0.15;
S0 = 100;
T = 1;
pasos = 100;
dt = T/pasos;
\[Kappa] = Exp\[Mu]J + \[Sigma]J^2/2] - 1;

(*Inicio trayectorias*)
traymodel = Table[0, {pasos + 1}];
traynormal = Table[0, {pasos + 1}];
traymodel[[1]] = S0;
traynormal[[1]] = S0;

DW = RandomVariate[NormalDistribution[0, 1], pasos];
PE = RandomVariate[PoissonDistribution\[Lambda]*dt], pasos];
jumpPoints = Table[0, {pasos + 1}];

(*Simulacion*)
Do[diffusionIncrement = (\[Mu] - \[Lambda]*\[Kappa])*dt + \[Sigma]*
  Sqrt[dt]*DW[[i]];

  traynormal[[i + 1]] = traynormal[[i]]*(1 + diffusionIncrement);
```

```
jumpIncrement = 0;

If[PE[[i]] > 0,
  jumpSizes =
    RandomVariate[LogNormalDistribution[\[Mu]J, \[Sigma]J], PE[[i]
      ]]];
jumpIncrement = Product[jumpSizes[[j]], {j, 1, PE[[i]]}] - 1;
AppendTo[
  jumpPoints, {i*dt,
    traymodel[[i]]*(1 + diffusionIncrement + jumpIncrement)}];];
traymodel[[i + 1]] =
  traymodel[[i]]*(1 + diffusionIncrement + jumpIncrement);, {i, 1,
  pasos}];

timePoints = Table[t*dt, {t, 0, pasos}];
priceData = Transpose[{timePoints, traymodel}];
diffusionData = Transpose[{timePoints, traynormal}];

mertonPlot =
  ListLinePlot[{priceData, diffusionData},
    PlotLegends ->
      Placed[{"Modelo de Salto-Difusion",
        "Componente de Difusion"}, {0.2, 0.75}],
    AxesLabel -> {"Tiempo", "Precio"},
    PlotStyle -> {{Blue}, {Opacity[0.7], Red, Dashed}},
    PlotRange -> All, Frame -> True, GridLines -> Automatic,
    ImageSize -> Large];
```

Modelo CIR

```
(*Parametros*)
a = 0.5;
```

```
\[Theta] = 1.0;
b = 0.5;
\[Sigma] = 2*Sqrt[a*b];
r0 = 0.21;
T = 15;
pasos = 120;
dt = T/pasos;
numtrays = 100000;

(*Inicio tablas de trayectorias*)
milsteintrays = Table[0., {numtrays}, {pasos + 1}];
eulertrays = Table[0., {numtrays}, {pasos + 1}];
hmtrays = Table[0., {numtrays}, {pasos + 1}];
drifttrays = Table[0., {numtrays}, {pasos + 1}];
e0trays = Table[0., {numtrays}, {pasos + 1}];
exacttrays = Table[0., {numtrays}, {pasos + 1}];

(*Simulacion*)
Do[
  milstein = Table[0., {pasos + 1}];
  euler = Table[0., {pasos + 1}];
  hm = Table[0., {pasos + 1}];
  drift = Table[0., {pasos + 1}];
  e0 = Table[0., {pasos + 1}];
  exact = Table[0., {pasos + 1}];
  milstein[[1]] = r0;
  euler[[1]] = r0;
  hm[[1]] = r0;
  drift[[1]] = r0;
  e0[[1]] = r0;
  exact[[1]] = r0;
```

```

Do[
  W = RandomVariate[NormalDistribution[0, 1]];
  (*Solucion exacta*)
  exact[[i + 1]] =
    Exp[-a*dt]*(Sqrt[exact[[i]]] +
      Sqrt[(b*(Exp[a*dt] - 1))/dt]*Sqrt[dt]*W)^2;
  (*\[Theta]-Milstein*)
  milstein[[
    i + 1]] = (1/(1 + a\[Theta]*dt))*(milstein[[i]] -
      a*dt*milstein[[i]] + a\[Theta]*dt*milstein[[i]] +
      a*b*dt + \[Sigma]*Sqrt[milstein[[i]]]*Sqrt[dt]*
      W + \[Sigma]^2/4 *((Sqrt[dt]*W)^2 - dt));
  (*Drift Implicit*)
  If[\[Sigma]^2 <= 4 a b,
    termSqrt = (\[Sigma]/2*Sqrt[dt]*W + Sqrt[drift[[i]]])^2 +
      4*(1 + a*dt/2)*(a*b - \[Sigma]^2/4)/2*dt;
    drift[[
      i + 1]] = ((\[Sigma]/2*Sqrt[dt]*W + Sqrt[drift[[i]]] +
      Sqrt[termSqrt])/(2*(1 + a*dt/2)))^2];
  (*E(0)*)
  If[a*dt != 2,
    e0[[i +
      1]] = ((1 - a*dt/2)*Sqrt[e0[[i]]] + \[Sigma]*Sqrt[dt]*
      W/(2*(1 - a*dt/2)))^2 + (a*b - \[Sigma]^2/4)*dt];
  (*Euler-Maruyama truncado*)
  currenteuler = Max[euler[[i]], 0];
  euler[[i + 1]] =
    euler[[i]] +
      a*(b - euler[[i]])*dt + \[Sigma]*Sqrt[currenteuler]*Sqrt[dt]*W;
  (*Higham-Mao*)
  currenthm = Abs[hm[[i]]];
  hm[[i + 1]] =

```

```
hm[[i]] +
  a*(b - hm[[i]])*dt + \[Sigma]*Sqrt[currenthm]*Sqrt[dt]*W, {i,
    1,
    pasos}];
milsteintrays[[tray]] = milstein;
drifttrays[[tray]] = drift;
e0trays[[tray]] = e0;
eulertrays[[tray]] = euler;
hmtrays[[tray]] = hm;
exacttrays[[tray]] = exact;
,
{tray, 1, numtrays}];

timePoints = Table[t*dt, {t, 0, pasos}];
(*Trayectoria media*)
meanMilsteintray = Mean[milsteintrays];
meanEulertray = Mean[eulertrays];
meanHmtray = Mean[hmtrays];
meanDrifttray = Mean[drifttrays];
meanE0tray = Mean[e0trays];
meanExacttray = Mean[exacttrays];
(*Distancia a la media*)
distanceMilstein = Abs[meanMilsteintray - b];
distanceEuler = Abs[meanEulertray - b];
distanceHm = Abs[meanHmtray - b];
distanceDrift = Abs[meanDrifttray - b];
distanceE0 = Abs[meanE0tray - b];
distanceExact = Abs[meanExacttray - b];

meanDataMilstein = Transpose[{timePoints, meanMilsteintray}];
meanDataEuler = Transpose[{timePoints, meanEulertray}];
meanDataHm = Transpose[{timePoints, meanHmtray}];
```

```
meanDataDrift = Transpose[{timePoints, meanDrifttray}];
meanDataE0 = Transpose[{timePoints, meanE0tray}];
meanDataExact = Transpose[{timePoints, meanExacttray}];

distanceDataMilstein = Transpose[{timePoints, distanceMilstein}];
distanceDataEuler = Transpose[{timePoints, distanceEuler}];
distanceDataHm = Transpose[{timePoints, distanceHm}];
distanceDataDrift = Transpose[{timePoints, distanceDrift}];
distanceDataE0 = Transpose[{timePoints, distanceE0}];
distanceDataExact = Transpose[{timePoints, distanceExact}];

(*grafico de la media*)
meanPlot =
  ListLinePlot[{meanDataExact, meanDataMilstein, meanDataEuler,
    meanDataHm, meanDataDrift, meanDataE0},
    PlotLabel -> "Trayectoria_ media",
    AxesLabel -> {"Tiempo_(t)", "Media"},
    PlotStyle -> {{Black, Thickness[0.003]}, {Red,
      Thickness[0.003]}, {Blue, Thickness[0.003]}, {Brown,
      Thickness[0.003]}, {Green, Thickness[0.003]}, {Orange,
      Thickness[0.003]}},
    PlotLegends ->
      Placed[{"Solucion_ exacta", "\[Theta]-Milstein",
        "Euler-Maruyama_truncado", "Higham-Mao", "Drift_Implicito",
        "E(0)"}, {1, 0.5}], PlotRange -> All, Frame -> True,
    GridLines -> Automatic, ImageSize -> Large, AspectRatio -> 1/2 ];

(*grafico de la distancia*)
distancePlot =
  ListLinePlot[{distanceDataMilstein, distanceDataEuler,
    distanceDataHm, distanceDataDrift, distanceDataE0},
    PlotLabel -> "Distancia_a_la media",
    AxesLabel -> {"Tiempo_(t)", "Distancia"},
    PlotStyle -> {{Red, Thickness[0.003]}, {Blue,
```

```
Thickness[0.003]], {Brown, Thickness[0.003]], {Green,
Thickness[0.003]], {Orange, Thickness[0.003]] },
PlotLegends ->
Placed[{"\[Theta]-Milstein", "Euler-Maruyama_\[truncado]",
"Higham-Mao", "Drift_\[Implicito]", "E(0)"}, {1, 0.5}],
PlotRange -> All, Frame -> True, GridLines -> Automatic,
ImageSize -> Large, AspectRatio -> 1/2, ScalingFunctions -> "Log"
];
```

Modelo de Heston

```
(*Parametros*)
\[Mu] = 0.1;
\[Kappa] = 2.0;
\[Theta] = 0.04;
\[Sigma]v = 0.3;
\[Rho] = -0.7;
S0 = 100;
v0 = 0.05;
T = 1;
pasos = 100;
dt = T/pasos;

(*Inicio de las trayectorias*)
price = Table[0, {pasos + 1}];
variance = Table[0, {pasos + 1}];
price[[1]] = S0;
variance[[1]] = v0;

(*Simulacion*)
Do[(*Incrementos brownianos correlacionados*)
Z1 = RandomVariate[NormalDistribution[0, 1]];
```

```
Z2 = \[Rho]*Z1 +
  Sqrt[1 - \[Rho]^2]*RandomVariate[NormalDistribution[0, 1]];
(*Euler-Maruyama truncado*)
currentVariance = Max[variance[[i]], 0];
variance[[i + 1]] =
  currentVariance + \[Kappa]*(\[Theta] - currentVariance)*
  dt + \[Sigma]v*Sqrt[currentVariance]*Sqrt[dt]*Z2;
price[[i + 1]] =
  price[[i]]*(1 + \[Mu]*dt + Sqrt[currentVariance]*Sqrt[dt]*Z1);, {
    i,
    1, pasos}];

timePoints = Table[t*dt, {t, 0, pasos}];
priceData = Transpose[{timePoints, price}];
varianceData = Transpose[{timePoints, variance}];

(*grafico del precio del activo*)
pricePlot =
  ListLinePlot[priceData, PlotLabel -> "Trayectoria del Precio",
    PlotStyle -> {Thickness[0.003], Blue}, PlotRange -> All,
    Frame -> True, FrameLabel -> {"Tiempo", "Precio"},
    GridLines -> Automatic, ImageSize -> Large];

(*grafico de la varianza*)
variancePlot =
  ListLinePlot[varianceData, PlotLabel -> "Varianza Instantanea",
    PlotStyle -> {Thickness[0.003], Red}, PlotRange -> All,
    Frame -> True, FrameLabel -> {"Tiempo", "Varianza"},
    GridLines -> Automatic, ImageSize -> Large];
```