



Universidade de Vigo

Trabajo Fin de Máster

Ecuaciones Diferenciales Estocásticas aplicadas a las Finanzas

Raúl J. Casas Fernández

Máster en Técnicas Estadísticas

Curso 2018-2019

Propuesta de Trabajo Fin de Máster

Título en galego: Ecuacións Diferenciais Estocásticas aplicadas ás Finanzas
Título en español: Ecuaciones Diferenciales Estocásticas aplicadas a las Finanzas
English title: Stochastic Differential Equations applied to Finance
Modalidad: Modalidad A
Autor/a: Raúl J. Casas Fernández, Universidad de Santiago de Compostela
Director/a: César A. Sánchez Sello, Universidad de Santiago de Compostela
Breve resumen del trabajo: Principales modelos paramétricos de ecuaciones diferenciales estocásticas y su aplicabilidad en un contexto económico-financiero.

Don César A. Sánchez Sello, Titular de Universidad de la Universidad de Santiago de Compostela informa que el Trabajo Fin de Máster titulado

Ecuaciones Diferenciales Estocásticas aplicadas a las Finanzas

fue realizado bajo su dirección por don Raúl J. Casas Fernández para el Máster en Técnicas Estadísticas. Estimando que el trabajo está terminado, dan su conformidad para su presentación y defensa ante un tribunal.

En Santiago de Compostela, a 8 de julio de 2019.

El director:

El autor:

Don César A. Sánchez Sello

Don Raúl J. Casas Fernández

Índice general

Resumen	I
1. Conceptos Básicos	1
1.1. Proceso de Wiener	1
1.2. Integral y Diferencial Estocástica	3
1.3. Procesos de Difusión	9
2. Ecuaciones diferenciales Estocásticas y sus Soluciones	11
2.1. Proceso de Ornstein-Uhlenbeck	11
2.2. Modelo del Movimiento Geométrico Browniano	16
2.3. Modelo de Cox-Ingersoll-Ross	19
2.4. Otros Modelos Paramétricos	24
3. Métodos de Simulación de Ecuaciones Diferenciales Estocásticas	29
3.1. Métodos de Discretización de las Ecuaciones Diferenciales Estocásticas	30
3.1.1. Esquema de Euler	31
3.1.2. Esquema de Milstein: $\beta, \gamma = 1$	37
3.1.3. Método de Predictor-Corrector	43
3.1.4. Segundo Esquema de Milstein	44
3.2. Métodos de Aproximación Local	45
3.2.1. Método Ozaki	46
3.2.2. Método Shoji-Ozaki	48
3.3. Métodos de Estimación Paramétrica	53
3.3.1. Métodos de Estimación de Verosimilitud Exacta	55
3.3.2. Métodos de Pseudo-Verosimilitud	64
3.3.3. Métodos de Aproximación Directa de la Verosimilitud	71

4. Ecuaciones Diferenciales Estocásticas y los Tipos de Interés	73
4.1. Modelo de Ait-Sahalia	73
4.1.1. Modelo de Vasicek o de Ornstein-Uhlenbeck	76
4.1.2. Modelo de Cox-Ingersoll-Ross (CIR)	99
4.1.3. Modelo CIR#	111
4.2. Conclusión	135

Resumen

Para modelar satisfactoriamente la dinámica de las variables económico-financieras como los tipos de interés, los precios de un activo o de un derivado, etc., se requiere del uso de la teoría de procesos estocásticos y, más concretamente, del empleo del cálculo estocástico que es de gran utilidad en el modelado en tiempo continuo. En los tiempos actuales, la preponderancia alcanzada por el cálculo de Itô en el ámbito económico-financiero ha sido realmente importante. Uno de los resultados cruciales del cálculo estocástico es el lema de Itô que se usa fundamentalmente para obtener diferenciales estocásticas; sensu stricto, el objeto de estudio del cálculo de Itô es la integral estocástica cuya notación simplificada, la ecuación diferencial estocástica, es el elemento sobre el que pivota este trabajo. Nuestro estudio únicamente se circunscribe al análisis de procesos de difusión que son solución a ecuaciones diferenciales estocásticas que contienen una componente determinista y una aleatoria o estocástica que involucra al proceso de Wiener o movimiento Browniano.

Entre los más destacables procesos estocásticos que son estudiados más detalladamente, están el movimiento geométrico Browniano o más conocido como el modelo de Black-Scholes-Merton (1973) cuya irrupción aportó un modelo para valorar una opción de compra europea sobre un activo subyacente cuya dinámica es conducida por un movimiento geométrico Browniano. Otro de los procesos de Pearson más relevantes objetivo de nuestro análisis es el proceso de Ornstein-Uhlenbeck o también conocido como proceso de Vasicek cuya contribución, cuatro años más tarde, está íntimamente ligada a los tipos de interés, concretamente a la tasa corta con reversión a la media que sigue una distribución normal y, por ende, alcanza valores positivos y negativos; esta última característica suponía una gran debilidad para su empleabilidad en la práctica financiera. Un modelo de aparición posterior surgido como una mejora del proceso de Vasicek es el proceso de Cox-Ingersoll-Ross cuya contribución más significativa descansa en que el tipo de interés «spot» sigue una distribución Chi cuadrada no centralizada, por lo que únicamente modela valores estrictamente positivos. Modelos paramétricos posteriores, casi todas particularidades del modelo de Aït-Sahalia, aportaron más complejidad o sofisticación al modelado de los tipos de interés.

Algunos de los modelos paramétricos arriba mencionados, poseen una solución explícita

única a una ecuación diferencial estocástica; asimismo, motivado por el cumplimiento de algunas asunciones (crecimiento lineal, global de Lipschitz, etc.) y de algunas restricciones en sus parámetros, además de markovianos son ergódicos lo cual nos regalan unas densidades de transición o condicional y estacionaria conocidas. En estos casos particulares, se analizan en profundidad los métodos de inferencia de verosimilitud exacta, los cuales, con algunas asunciones adicionales, nos aseguran unos estimadores consistentes y asintóticamente normales. En los casos restantes, se exponen los métodos de pseudo-verosimilitud que no aproximan la densidad condicional directamente, sino que la trayectoria del proceso de forma que su versión discretizada por Euler, Milstein, Ozaki, Shoji-Ozaki o cualquier otra, tenga una verosimilitud usable para la inferencia paramétrica.

Por último, se procede a mimetizar la operativa de un presunto experto financiero cuyo interés principal se centra en el modelado de los tipos de interés. Para ello, hacemos uso de una muestra de 2548 observaciones de tipos de interés diarios en el mercado europeo (EONIA) que, va desde febrero de 2009 hasta febrero de 2019, y, mediante el proceso de Ornstein-Uhlenbeck y el proceso CIR, se estudia cómo se ajustan estos modelos a los datos de nuestra muestra. Ante los previsible obstáculos que nos podríamos encontrar, proponemos una serie de metodologías que contribuyen efectivamente a la aplicabilidad de estos modelos, o de sus derivados, en el mundo financiero.

Abstract

To effectively model the dynamic of economic-financial variables such as the interest rates, the price of an active or a derivative, etc, it requires the use of the theory of stochastic processes and, more specifically, the use of stochastic calculus which is the great utility in continuous time modelling. Lately, the relevance reached by Itô's calculus in the economic-finance context, has been significantly important. One of the key results attained in the stochastic calculus is the Itô's lemma that its main use consists in obtaining stochastic differentials; sensu stricto, the aim of studying the Itô's calculus is the stochastic integral whose simplified notation, the

stochastic differential equation, is the pivotal element of this work.

Among those most outstanding stochastic processes which were analyzed in more detail are the geometric Brownian motion, or also known as the Black-Scholes-Merton model (1973); its irruption supposed a model to value a European call option on an underlying whose dynamic is guided by a geometric Brownian motion. Another Pearson's process is the well known Ornstein-Uhlenbeck process or also named Vasicek's process which main contribution, four years later than O-U's appearance, is closely linked to the interest rates, or more specifically, to the mean reverting short rate, that follows a normal distribution and, thus, may reach positive and negative values; this last feature, in those times, was considered a great weakness in terms of applicability to the financial field. Years later, as a presumable improvement to the Vasicek's model, is the Cox-Ingersoll-Ross's model which main contribution rests on the findings that the spot interest rate follows a non-central Chi-squared, therefore it models strictly positive values. Other parametric models, most of them are particular cases of the Ait-Sahalia model, brought even more complexity into the interest rate modelling.

Some of the later parametric models, have a unique explicit solution to a stochastic differential equation; furthermore, if some assumptions (i.e. linear growth assumption, global Lipschitz assumption, etc.) and some parametric restrictions are accomplished, then the processes not only are Markovian but also ergodic, and subsequently their transitional or conditional and stationary density are known. In these particular cases, we deeply analyze the exact likelihood inference methods, that, with additional assumptions, consistent and asymptotically efficient estimators may be obtained. In the remaining cases, the pseudo-likelihood methods that approach the conditional density indirectly are dealt, instead they simulate the discretized version by the Euler, Milstein, Ozaki, Shoji-Ozaki, etc, to make parametric inference.

Finally, we try to imitate the procedure of a financial practitioner whose main target points at the interest rate modelling. To do it, we make use of a sample containing 2548 observations of European daily interest rates (EONIA). Its sampling period starts at February 23rd, in 2009 and finishes at February 20th, in 2019, and through a Ornstein-Uhlenbeck and CIR processes, we study how well these models adjust to our sample data. Throughout our analysis, we propose a complete set of methodologies, in order to overcome foreseeable obstacles, that effectively contribute to the applicability of these models, or their derived models, to the financial world.

Capítulo 1

Conceptos Básicos

A lo largo del documento, se hará uso de diversos conceptos que, o bien constituyen ingredientes básicos de los modelos que describen una dinámica estocástica, o bien forman parte de la solución a las ecuaciones diferenciales estocásticas. Así, en la siguiente sección, brevemente se revisará en qué consiste un proceso de Wiener; seguidamente, se afrontará la exposición de los conceptos de integral estocástica y ecuación diferencial estocástica y, finalmente, se mostrará en qué consiste un proceso de difusión.

1.1. Proceso de Wiener

El proceso de Wiener o también llamado movimiento Browniano, denotado como $W = W(t), t \geq 0$, se podría definir de múltiples formas, no obstante, la elegida por nosotros es la que sigue: es un proceso Gaussiano con trayectorias continuas y con incrementos independientes, de tal manera que

$$W(0) = 0, \tag{1.1}$$

con probabilidad 1

$$\mathbb{E}W(t) = 0 \tag{1.2}$$

para todo $t > 0$

y, para todo $0 \leq s \leq t$,

$$\text{Var}(W(t) - W(s)) = t - s \tag{1.3}$$

Los incrementos siguen una distribución normal de media, 0 y de varianza, $t - s$, para todo $0 \leq s \leq t$ y son independientes, esto es, dados dos cualesquiera intervalos disjuntos, digamos (t_1, t_2) , (t_3, t_4) , con $t_1 \leq t_2 \leq t_3 \leq t_4$, sus respectivos incrementos, es decir, $(W(t_2) - W(t_1))$ y

$(W(t_4) - W(t_3))$ son independientes.

Así, por la característica de los incrementos anteriormente descrita, el proceso de Wiener tiene una variación simple infinita, $V_t(W) = \infty$, y una covarianza, $Cov(W(s), W(t)) = t \wedge s = \min(s, t)$. El proceso de Wiener posee unas trayectorias continuas pero no diferenciables.

Debemos comentar que el proceso de Wiener, $W(t)$, tiene trayectorias que pertenecen al espacio de funciones de $[0, T]$ a \mathbb{R} , $L^2([0, T])$, definido como

$$L^2 = \{f : [0, T] \rightarrow \mathbb{R} : \|f\|_2 < \infty\},$$

donde

$$\|f\|_2 = \left(\int_0^T |f(t)|^2 dt \right)^{\frac{1}{2}}.$$

También es conveniente definir la expansión de Karhunen-Loève del proceso de Wiener, W , como una expansión L^2 de procesos aleatorios en términos de una secuencia de variables independientes funciones ortogonales y coeficientes aleatorios como el proceso de Wiener, que tiene una colección de variables Gaussianas independientes. Su utilidad radica en su capacidad de representación de un proceso en un intervalo fijo de tiempo, i.e. $[0, T]$. La expansión de Karhunen-Loève de un proceso de Wiener sería como sigue:

$$W(t) = W(t, w) = \sum_{i=0}^{\infty} Z_i(w) \phi_i(t), 0 \leq t \leq T,$$

donde

$$\phi_i(t) = \frac{2\sqrt{2T}}{(2i+1)\pi} \sin\left(\frac{(2i+1)\pi}{2T}t\right),$$

que es la función phi definida en el código.

Las funciones ϕ_i conforman una base de funciones ortogonales y las Z_i son una secuencia de variables aleatorias Gaussianas estándar, independientes e idénticamente distribuidas.

1.2. Integral y Diferencial Estocástica

En el contexto matemático-financiero el cálculo estocástico es una de las herramientas más usadas. Mucha de la teoría económico-financiera se apoya en el cálculo de Itô. A continuación, se desarrolla el concepto de integral estocástica que es el objeto de estudio del cálculo estocástico.

Sea $W(t), t \geq 0$ un movimiento Browniano definido sobre un espacio fijo de probabilidad con filtración aumentada $(\Omega, \mathfrak{F}, (F_t)_{t \geq 0}, \mathbb{P})$.

Consideremos una ecuación del tipo

$$S_t = S_0 + \int_0^t \mu(S_u, u) du + \int_0^t \sigma(S_u, u) dW_u. \quad (1.4)$$

así, podemos presentar la integral estocástica con respecto a un movimiento Browniano:

$$I(X) = \int_0^T X(u) dW(u)$$

La integral estocástica de un proceso se define como Sea $\{X(t), 0 \leq t \leq T\}$, un proceso estocástico adaptado a la filtración generada por un movimiento Browniano, tal que $\int_0^T \mathbb{E}X(x)^2 ds < +\infty$.

$$I_t(X) = \int_0^t X_s dW_s = \lim_{\|\Pi_n\| \rightarrow 0} \sum_{i=0}^{n-1} X(t_i)(W(t_{i+1}) - W(t_i)), \quad (1.5)$$

donde la convergencia es en media cuadrática y $t_i \in \Pi_n$; siendo Π_n , con $n \geq 1$ una sucesión de particiones $0 = t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n = T$.

También se podría definir de esta otra forma:

$$I_t(X) = \int_0^t X_s dW_s,$$

tal que

$$\lim_{n \rightarrow \infty} \mathbb{E}^{\mathbb{P}} \left[\sum_{i=0}^{n-1} X(t_i)(W(t_{i+1}) - W(t_i)) - I_t(X) \right]^2 = 0,$$

donde se observa que la convergencia es en media cuadrática, en \mathfrak{L}^2 , donde $\{W_t, t \geq 0\}$ es

un movimiento Browniano estándar y donde $0 = t_0 < t_1 < t_2 < \dots < t_n = t$ es una partición del intervalo de tiempo $[0, t]$, tal que $t_i - t_{i-1} = t/n$, para $i = 1, 2, \dots, n$.

Se podría definir la integral estocástica como el límite en media cuadrática de un sumatorio de términos que comprenden incrementos independientes de un movimiento Browniano.

En la coyuntura económico-financiera, una notación diferente de 1.4 es una ecuación diferencial estocástica de la forma

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t. \quad (1.6)$$

¿Cómo se puede presentar la heurística que hay detrás de una EDE, para modelar el comportamiento del precio de un activo financiero, por ejemplo? Se va a describir, a través de un ejemplo fácil de entender, su componente determinista y su componente estocástica que capture esos movimientos diarios o perturbaciones aleatorias que se producen en el precio del activo. Supongamos una población, una granja de parejas de cobayas de laboratorio, que se reproducen geoméricamente en tiempo discreto; supongamos que su tasa de mortalidad es nula. Asimismo, supongamos que la tasa de reproducción es constante y la denotamos por μ , y supongamos que $\mu = 2$. En el tiempo, $t = 0$, el granjero tiene una única pareja de cobayas y lo denotaremos por $N_0 = 1$. En $t = 1$, la población, N_1 , será de tres, pues $N_1 = N_0 + \mu \cdot N_0$. En $t = 2$, la población, N_2 , ya será de nueve, ya que $N_2 = N_1 + \mu \cdot N_1$ y así, sucesivamente. Por ende, se deriva que, la población de cobayas en tiempo, t , vendrá determinada por:

$$N_t = N_{t-1} + \mu \cdot N_{t-1} = N_{t-1}(1 + \mu) = N_0(1 + \mu)^t.$$

Análogamente, si en lugar de una población inicial de cobayas de N_0 , se dispone de una inversión inicial, S_0 , que produce una rentabilidad geométrica a una tasa anualizada, μ , durante un período de t años, tendríamos:

$$S_t = S_{t-1}(1 + \mu) = S_0(1 + \mu)^t;$$

y si en lugar de una tasa anualizada, la suma invertida crece a una tasa compuesta n veces, la expresión anterior sería:

$$S_t = S_0\left(1 + \frac{\mu}{n}\right)^{nt}.$$

Si fijamos t , y $nt \rightarrow \infty$, entonces

$$S_t = S_0 \left(1 + \frac{\mu \cdot t}{n \cdot t}\right)^{nt}.$$

Así, si $e^\mu \approx 1 + \mu$, para $\mu \rightarrow 0$, tendríamos

$$S_t \approx S_0 \cdot e^{\mu t}. \quad (1.7)$$

Por lo que se pasa de un crecimiento geométrico a un incremento exponencial, y se pasa de un tiempo discreto a un tiempo continuo, permaneciendo constante la tasa de crecimiento, μ . Con ello, finalmente tenemos que la ecuación última resulta ser solución de la siguiente ecuación diferencial de primer orden

$$\frac{dS_t}{dt} = \mu \cdot S_t,$$

que se puede reescribir como

$$\mu = \frac{dS_t}{S_t} \frac{1}{dt}.$$

Lo que indica que la tasa μ , constante, es igual a la tasa de crecimiento del precio del activo por unidad de tiempo. Para dotar de más credibilidad al relato, le añadimos una serie de perturbaciones aleatorias que siguen un movimiento Browniano estandarizado, $dW_t \sim \mathbb{N}(0, dt)$. Comentar que estas perturbaciones representan variaciones positivas o negativas en los precios, así, el proceso quedaría como

$$\frac{dS_t}{S_t} = \mu dt + dW_t. \quad (1.8)$$

Hemos hablado del signo de estas perturbaciones, positivas o negativas, pero no así de su magnitud, grande o pequeña. Si introducimos el parámetro σ en la ecuación 1.8, la dinámica del proceso quedaría definida mediante la ecuación que sigue:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,$$

con $\sigma > 0$. Así si los valores de σ son altos, tendríamos perturbaciones importantes y si, por el contrario, los valores son bajos, la magnitud de las perturbaciones será pequeña; proceso

que, aparentemente, capta de una manera obviamente simplista lo que sucede en la realidad. Por todo lo anterior, la variación de los precios de los activos estaría conformada por una componente determinista o de tendencia, y una componente estocástica o de difusión o aleatoria. Ya estamos en disposición de formular con más ortodoxia matemático-financiera el proceso explicado anteriormente. Sea $S(t)$, $t \geq 0$, valor de un activo en tiempo t . Sea $\frac{\Delta S}{S}$, el rendimiento del activo.

$$\frac{\Delta S}{S} = \mu \Delta t + \sigma \Delta W,$$

donde Δt , representa la variación en el intervalo de tiempo $[t, t + \Delta t]$, Δ representa la variación del ruido o innovaciones o shocks, $\mu \Delta t$ representa la aportación determinista (relacionada con la tasa de interés del activo libre de riesgo), y $\sigma \Delta X$, la aportación estocástica (relacionada con las innovaciones o shocks y con la volatilidad). Se asume que el ruido es Gaussiano, $\Delta X \approx \mathbb{N}(0, 1)$. Si asumimos que las innovaciones son independientes, entonces, X es un proceso de Wiener. Así pues, se tendría que

$$\frac{\Delta S}{S} = \mu \Delta t + \sigma \Delta W.$$

Si $\Delta t \rightarrow 0$, se obtendría la ecuación diferencial estocástica siguiente:

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t). \quad (1.9)$$

Matemáticamente, debido a que la variación de un proceso de Wiener, $dW(t)$, es no finita, y, como ya hemos visto en la sección anterior, el proceso de Wiener tiene trayectorias continuas y no diferenciables, la expresión anterior carece de significado. Para ello, se reescribe como sigue

$$S(t) = S(0) + \mu \int_0^t S(u)du + \sigma \int_0^t S(u)dW(u), \quad (1.10)$$

A continuación, se destacan algunas propiedades de la integral de Itô que son de gran utilidad. Así, X es Itô integrable si cumple que:

$$\mathbb{E}\left(\int_0^T X(s)dW(s)\right) = 0$$

y

$$\text{Var}\left(\int_0^T X(s)dW(s)\right) = \int_0^T \mathbb{E}X^2(t)dt,$$

donde esta última propiedad podría escribirse como

$$\mathbb{E}\left(\int_0^T X(s)dW(s)\right)^2 = \int_0^T \mathbb{E}X^2(t)dt,$$

conocida como la isometría de Itô. Adicionalmente, debe poseer la propiedad de la linealidad. Dadas dos integrales de Itô, X e Y y, dadas dos constantes, $a, b \in \mathbb{R}$, entonces

$$\int_0^T (aX(t) + bY(t))dW(t) = a \int_0^T X(t)dW(t) + b \int_0^T Y(t)dW(t).$$

Asimismo, existe una tercera propiedad, la propiedad de martingala, que asevera que dada una integral de Itô, X , y, dada una constante, $M(0)$, entonces el proceso

$$M(t) = M(0) + \int_0^t X(s)dW(s), \quad (1.11)$$

es una martingala con $M(0)$ una constante.

Definimos lo que es un proceso de Itô. Un proceso de Itô es un proceso estocástico de la forma

$$X_t = X_0 + \int_0^t g(s)ds + \int_0^t h(s)dW_s. \quad (1.12)$$

Donde $g(\cdot, \cdot)$ y $h(\cdot, \cdot)$, son dos funciones aleatorias adaptadas y medibles progresivamente, tales que

$$\mathbb{P}\left\{\int_0^T |g(t, w)|dt < \infty\right\} = 1 \quad (1.13)$$

y

$$\mathbb{P}\left\{\int_0^T h(t, w)^2 dt < \infty\right\} = 1 \quad (1.14)$$

Un modo de simular trayectorias de algún proceso paramétrico (Ornstein-Uhlenbeck, Vasicek, Black-Scholes, Cox-Ingersoll-Ross), aunque hay otros como ya se verá en otros capítulos, es mediante la simulación de una integral de Itô; a continuación, se muestra un código donde

simularemos una integral estocástica.

```
W <- BM()
t <- time(W)
N <- length(t)
x <- 8
theta <- 6
sigma <- 3
X <- numeric(N)
X[1] <- x
ito.sum <- c(0, sapply(2:N, function(x){
  exp(-theta*(t[x]-t[x-1]))*(W[x]-W[x-1]))})
X <- sapply(1:N, function(x) {
  X[1]*exp(-theta*t[x])+sum(ito.sum[1:x])})
X <- ts(X, start = start(W), deltat = deltat(W))
plot(X, main = "Ornstein-Uhlenbeck")
```

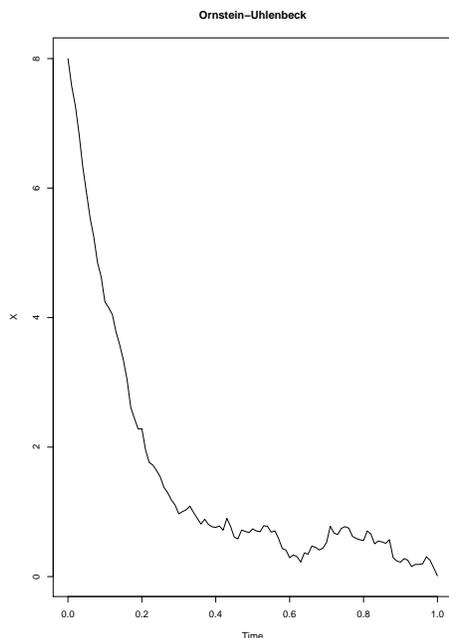


Figura 1.1: Simulación de una trayectoria de un proceso de Ornstein-Uhlenbeck mediante la aproximación de una integral de Itô.

Básicamente, se trata de una simulación de sumas de Itô como un modo de simular trayectorias mediante la simulación de la integral estocástica. Recordemos que la integral estocástica definida en 1.5 contiene una suma de Itô que va desde cero hasta $n - 1$.

1.3. Procesos de Difusión

Los procesos de difusión engloban la mayor parte de los procesos que son solución a una ecuación diferencial estocástica de la forma

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t), \quad (1.15)$$

donde $b(\cdot, \cdot)$ es el coeficiente drift y donde $\sigma(\cdot, \cdot)$ es el coeficiente difusión de 1.15; con $X(0)$ como condición inicial, que podría ser aleatoria o no; en caso de que sea aleatoria, sería independiente de la σ -álgebra generada por W y satisfaría la condición $\mathbb{E}|Z|^2 < \infty$, donde

$Z = X(0)$, sería la condición inicial aleatoria.

La ecuación 1.15, debe ser pensada como una integral de Itô de la forma

$$X(t) = X(0) + \int_0^T b(u, X(u))du + \int_0^T \sigma(u, X(u))dW(u)$$

y dado que en un proceso de Itô, debe cumplir dos asunciones si se desea tener una solución fuerte, única, continua y adaptada, tal que

$$\mathbb{E}\left\{\int_0^T |X_t|^2\right\} < \infty.$$

Las dos condiciones para obtener una solución X del tipo fuerte son:

- 1.- La condición Global de Lipschitz: $\forall x, y \in \mathbb{R}$ y $t \in [0, T]$, entonces $\exists K < +\infty$, tal que

$$|(b(t, x) - b(t, y))| + |(\sigma(t, x) - \sigma(t, y))| < K|x - y|, \quad (1.16)$$

con K , constante.

- 2.- La condición de crecimiento lineal: $\forall x, y \in \mathbb{R}$ y $t \in [0, T]$, entonces $\exists C < +\infty$, tal que

$$|(b(t, x))| + |(\sigma(t, x))| < C(1 + |x|), \quad (1.17)$$

con C , constante.

Esta solución X es lo que se conoce como un proceso de difusión que, si es del tipo fuerte, implica unicidad y, por tanto, la existencia de una única trayectoria del resultado. La principal diferencia con un proceso de difusión del tipo débil solución a una ecuación diferencia estocástica, estriba en que, una vez asumidas otras condiciones, dadas cualesquiera dos soluciones débiles, éstas no poseen idénticas trayectorias, pero, sí que poseen idénticas distribuciones.

$$P\left\{\int_0^T |g(t, w)|dt < \infty\right\} = 1$$

y

$$P\left\{\int_0^T h(t, w)^2 dt < \infty\right\} = 1$$

Capítulo 2

Ecuaciones diferenciales Estocásticas y sus Soluciones

En este capítulo describiremos los procesos estocásticos soluciones, únicas o no, de ecuaciones diferenciales estocásticas. Asimismo, se expondrán sus principales características y se analizará cada proceso en función de su densidad estacionaria, de su distribución condicional, de sus momentos y de su función de covarianza.

En la sección siguiente trataremos con el movimiento geométrico Browniano o modelo de Black-Scholes-Merton cuya aparición marcó un antes y después en las matemáticas financieras empleadas para la valuación de productos derivados cuyos subyacentes conforman un extenso conjunto de productos financieros (acciones, tipos de cambio, etc.); posteriormente estudiaremos el proceso de Ornstein-Uhlenbeck o también conocido como proceso de Vasicek, cuya aportación más relevante se circunscribe a los tipos de interés. En la sección tercera, nuestros esfuerzos irán destinados a exponer con detalle uno de los procesos estocásticos más conocidos y difundidos, el modelo de Cox-Ingersoll o Ross; todos ellos procesos de Pearson casos particulares de la familia de modelos de Chan-Karolyi-Longstaff-Sanders y ésta, a su vez, también se trata de un caso concreto del modelo de Aït-Sahalia. Finalmente, concluimos con un breve recorrido por otros modelos paramétricos con especial relevancia en el contexto financiero.

2.1. Proceso de Ornstein-Uhlenbeck

Sea $\{W_t, t \geq 0\}$, un movimiento Browniano definido sobre un espacio de probabilidad $(\Omega, \mathfrak{F}, \mathbb{P})$ y sea $\mathbb{F} = \{\mathfrak{F}_t, t \geq 0\}$, su filtración aumentada. Supongamos que la dinámica estocástica de

X_t es conducida por la ecuación diferencial estocástica siguiente

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \quad (2.1)$$

con la condición inicial no aleatoria, $X_0 = x_0$, con $\theta, \mu \in \mathbb{R}$ y $\sigma \in \mathbb{R}_+$.

El proceso de Ornstein-Uhlenbeck es la única solución a la ecuación 2.1, esto es, se trata de una solución de tipo fuerte. Esta unicidad está inextricablemente unida al cumplimiento de la condición global de Lipschitz, así como de la condición de crecimiento lineal. Recordemos, a través de un ejemplo de una dinámica de tipos de interés de tasa corta, guiada por una ecuación diferencial estocástica similar a la 2.1, cómo nos aseguramos de que la ecuación diferencial estocástica¹ que sigue, tenga una solución única:

$$dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t,$$

con $\mu(t, r_t)$ y $\sigma(t, r_t)$ son procesos adaptados a la filtración $\mathbb{F} = \{\mathfrak{F}_t, t \geq 0\}$. La expresión genuinamente apropiada sería

$$r_t = r_0 + \int_0^t \mu(u, r_u)du + \int_0^t \sigma(u, r_u)dW_u.$$

Para que exista unicidad en la solución se requiere que tanto el parámetro drift, $\mu(t, r_t)$, como el parámetro difusión, $\sigma(t, r_t)$, cumplan las dos condiciones antes mencionadas. De modo que debe verificar:

$$|\mu(t, r) - \mu(t, s)| + |\sigma(t, r) - \sigma(t, s)| \leq K|r - s|.$$

Para todo $t \in [0, T]$ y para todo $r, s \in \mathbb{R}$; con $K < +\infty$, una constante independiente de r y s . Y, además, debe satisfacer que

$$\mu^2(t, r) + \sigma^2(t, r) \leq C(1 + r^2).$$

Para todo $t \in [0, T]$ y para todo $r \in \mathbb{R}$; con $C < +\infty$.

A modo de homenaje a Oldrich Alfons Vasicek, el proceso de Ornstein-Uhlenbeck también se le conoce como proceso de Vasicek, no obstante, s.s. éste estaría modelado de la siguiente

¹Tengamos siempre presente que, cuando se hable de ecuaciones diferenciales estocásticas, no es más que una notación simplificada, pues es, en realidad, una integral estocástica.

forma

$$dX_t = -\theta_2 X_t dt + \theta_3 dW_t, \quad X_0 = x_0,$$

que, usando la misma parametrización que hasta ahora, sería

$$dX_t = -\theta X_t dt + \sigma dW_t, \quad X_0 = x_0,$$

donde $\theta_2 = \theta$ y $\theta_3 = \sigma$. Por tanto, en este proceso $\theta_1 = \mu = 0$.

Una propiedad interesante del proceso de Ornstein-Uhlenbeck es que, para $\theta > 0$, se trataría de un proceso que revierte a la media, lo que significa que el proceso tiende a oscilar alrededor de un estado de equilibrio que es μ . Otra propiedad destacable del proceso es que se trata de un proceso con varianza finita, contrariamente al proceso de Wiener.

$$\mathbb{E}\left\{\int_0^T |X_t|^2 dt\right\} < \infty.$$

La parametrización de la ecuación diferencial estocástica 2.1 representa la forma del proceso de Ornstein-Uhlenbeck más ampliamente usada en un ámbito económico-financiero, donde σ es la volatilidad, μ es la media en el largo plazo y θ es la velocidad de reversión a la media.

Sin embargo, en el ámbito matemático, el proceso de Ornstein-Uhlenbeck como solución única de una ecuación diferencial estocástica de la forma

$$dX_t = (\theta_1 - \theta_2 X_t) dt + \theta_3 dW_t, \tag{2.2}$$

con la condición inicial no aleatoria, $X_0 = x_0$, con $\theta_1, \theta_2 \in \mathbb{R}$ y $\theta_3 \in \mathbb{R}_+$.

La solución explícita de la ecuación 2.2 sería:

$$X_t = \frac{\theta_1}{\theta_2} + \left(x_0 - \frac{\theta_1}{\theta_2}\right) e^{-\theta_2 t} + \theta_3 \int_0^t e^{-\theta_2(t-u)} dW_u. \tag{2.3}$$

A continuación, se expone cómo se alcanza esta solución de tipo fuerte.

Si $f(t, x) = x e^{\theta_2 t}$, tendríamos que sus derivadas primeras parciales y su derivada segunda respecto de x serían

$$f_t(t, x) = \theta_2 f(t, x),$$

$$f_x(t, x) = e^{(\theta_2 t)},$$

$$f_{xx}(t, x) = 0.$$

Por tanto,

$$\begin{aligned} X_t e^{\theta_2 t} &= f(0, X_0) + \int_0^t \theta_2 X_u e^{\theta_2 u} du + \int_0^t e^{\theta_2 u} dX_u \\ &= x_0 + \int_0^t \theta_2 X_u e^{\theta_2 u} du + \int_0^t e^{\theta_2 u} \{(\theta_1 - \theta_2 X_u) du + \theta_3 dW_u\} \\ &= x_0 + \frac{\theta_1}{\theta_2} e^{\theta_2 t} - 1 + \theta_3 \int_0^t e^{\theta_2 u} dW_u. \end{aligned}$$

Luego cumple la ecuación 2.3.

Con esta parametrización, ya hemos explicado que si $\theta_2 > 0$, el proceso oscila alrededor de un estado de equilibrio.

Que el parámetro θ_2 sea estrictamente positivo, nos obsequia con la característica de ergodicidad, y esta propiedad implica que

$$\mathbb{P}\left\{\frac{1}{T} \int_0^T h(X_t) dt \rightarrow \int_{-\infty}^{+\infty} h(x) \pi(x) dx\right\} = 1,$$

que sería igual a

$$\mathbb{P}\{\mathbb{E}h(\xi)\} = 1$$

para cualquier función medible $h(\cdot)$, donde $\pi(\cdot)$ resulta ser la densidad estacionaria del proceso y ξ es una variable aleatoria con función de densidad $\pi(\cdot)$.

La distribución incondicional del proceso sigue una distribución normal con las siguientes media, varianza y función de covarianza estacionarias:

$$m(X_t) = \frac{\theta_1}{\theta_2},$$

$$Var(X_t) = \frac{\theta_3^2}{2\theta_2},$$

$$\text{Cov}(X_t, X_s) = \frac{\theta_3^2 e^{-\theta_2|t-s|}}{2\theta_2}.$$

Su densidad condicional o de transición, esto es, la densidad de la distribución de X_t dada una condición inicial, aleatoria o no, X_0 , es una normal, con media, varianza y covarianza condicionales:

$$\mathbb{E}_\theta(X_t|X_0 = x_0) = \frac{\theta_1}{\theta_2} + \left(x_0 - \frac{\theta_1}{\theta_2}\right)e^{-\theta_2 t}, \quad (2.4)$$

$$\text{Var}_\theta(X_t|X_0 = x_0) = \frac{\theta_3^2(1 - e^{-2\theta_2 t})}{2\theta_2}, \quad (2.5)$$

$$\text{Cov}(X_s, X_t|X_0 = x_0) = \frac{\theta_3^2}{2\theta_2} e^{-\theta_2(s+t)} \left(e^{2\theta_2(s \wedge t)} - 1\right).$$

Como sabemos que la densidad de transición es Gaussiana, podemos simularla a través de su media y varianza condicional, así:

```
set.seed(1234)
x0 <- 0
theta <- c(2,1,0.25)
### --- Densidad Transitoria --- ###
dcOU <- function(t, x0, theta){
  c.mean <- theta[1]/theta[2]+(x0-theta[1]/theta[2]) * exp(-theta[2]*t) # Media Condicional
  c.var <- theta[3]^2*(1-exp(-2*theta[2]*t))/(2*theta[2]) # Varianza Condicional
  rnorm(1, mean = c.mean, sd = sqrt(c.var))
}
require(sde)
n <- 200
x <- c()
x[1]=dcOU(t=0.1,x0,theta)
for (i in 2:n){
  x[i]=dcOU(t=0.1,x[i-1],theta)
}
Xc <- ts(c(x0,x))
```

```
plot(Xc, main = "Simulación: Dens. Trans. Proceso OU")
abline(h=2,lty=2,col=2) # Media a Largo plazo
```

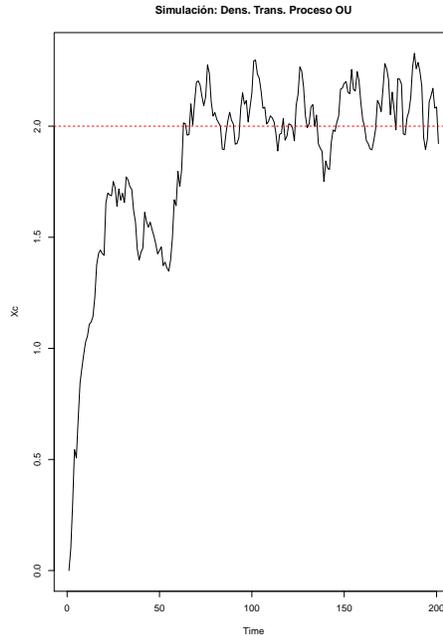


Figura 2.1: Densidad de transición o condicional simulada de un proceso de Ornstein-Uhlenbeck; donde se observa cómo el proceso a partir de un determinado t oscila en torno a la media en el largo plazo, $\theta_1 = 2$.

2.2. Modelo del Movimiento Geométrico Browniano

Un proceso que suele usarse con mucha frecuencia en el ámbito financiero al modelar la dinámica de los activos financieros (acciones, futuros, derivados, etc.) es el movimiento geométrico Browniano. Éste, en finanzas, se define como una función del movimiento Browniano de la forma

$$X(t) = x \exp \left\{ \left(r - \frac{\sigma^2}{2} \right) t + \sigma W(t) \right\}, \quad (2.6)$$

con $t > 0$, $x = X_0$, $x \in \mathbb{R}$ y donde $\{W(t), t \geq 0\}$ es un movimiento Browniano estándar y σ , es la volatilidad estrictamente positiva, y r es la tasa de interés libre de riesgo.

Si reparametrizamos esta ecuación 2.6, y denotamos por θ_1 en lugar de r , y θ_2 en vez de σ , nos quedaría así

$$X(t) = x \exp \left\{ \left(\theta_1 - \frac{\theta_2^2}{2} \right) t + \theta_2 W(t) \right\}, \quad (2.7)$$

con $t > 0$, $x = X_0$, $x \in \mathbb{R}$ y donde $\{W(t), t \geq 0\}$ es un movimiento Browniano estándar; análogamente, con $\theta_2 > 0$ y θ_1 una constante (la tasa libre de riesgo).

Este movimiento geométrico Browniano 2.7, es la solución explícita a la ecuación diferencial estocástica siguiente

$$dX_t = \theta_1 X_t dt + \theta_2 X_t dW_t,$$

con $\theta_2 > 0$ y $X_0 = x_0$.

Así como en el modelo de la sección última su función de densidad condicional era Gaussiana, en el caso del movimiento Browniano geométrico, su densidad de transición es log-normal. Para ver cómo se obtiene esta función de densidad, circunscribamos nuestro análisis en un contexto económico-financiero; el modelo del movimiento geométrico Browniano suele denominarse como el modelo de Black-Scholes-Merton, usado fundamentalmente para determinar los precios de los derivados (opciones, esencialmente). No es objetivo prioritario de nuestro análisis, limitarnos a una familia concreta de productos financieros, pero ello no es óbice como para usar a modo de ejemplo, la dinámica del precio de un activo financiero cualquiera.

Así, sea $(W_t)_{t \in [0, T]}$ un proceso de Wiener definido sobre un espacio fijo de probabilidad con una filtración $(\Omega, \mathfrak{F}, (\mathfrak{F}_t)_{t \in [0, T]}, \mathbb{P})$. Sea X_t , el precio de un activo en el tiempo t , que es conducido por el movimiento geométrico Browniano

$$dX_t = \theta_1 X_t dt + \theta_2 X_t dW_t,$$

donde el parámetro drift, $\theta_1 \in \mathbb{R}$, es el rendimiento medio esperado del activo en cuestión, y, el parámetro, $\theta_2 > 0$ es la volatilidad por unidad de tiempo. Así, al hacer uso del lema de Itô:

$$d(\ln X_t) = \left(\theta_1 - \frac{1}{2}\theta_2^2\right)dt + \theta_2 dW_t.$$

Si $\Delta t = T - t$, entonces

$$\ln X_T - \ln X_t = \left(\theta_1 - \frac{1}{2}\theta_2^2\right)(T - t) + \theta_2\sqrt{T - t}\epsilon,$$

donde $\epsilon \sim \mathbb{N}(0, 1)$, cuya función de densidad es $\phi(\epsilon) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\epsilon^2}$, con $\epsilon \in \mathbb{R}$. Así pues

$$\ln \frac{X_T}{X_t} \sim \mathbb{N}\left(\left(\theta_1 - \frac{1}{2}\theta_2^2\right)(T - t), \theta_2^2(T - t)\right).$$

Ahora denotamos g a la transformación

$$g(\epsilon) := X_T = X_t \exp\left\{\left(\theta_1 - \frac{\theta_2^2}{2}\right)(T - t) + \theta_2\sqrt{T - t}\epsilon\right\},$$

y por tanto

$$g^{-1}(X_T) = \frac{\ln\left(\frac{X_T}{X_t}\right) - \left(\theta_1 - \frac{1}{2}\theta_2^2\right)(T - t)}{\theta_2\sqrt{T - t}}.$$

Por ende, la función de densidad de X_T condicionada a X_t sería

$$f_{X_T|x_t}(x|X_t) = p_\theta(t, x|x_0) = \phi(g^{-1}(x))\left|\frac{dg^{-1}(x)}{dx}\right|,$$

donde ϕ es la función de densidad normal estándar.

El éxito del modelo de Black-Scholes-Merton inicialmente descansa sobre el extensivo uso de la siguiente función de densidad transitoria en el cálculo del valor esperado del valor intrínseco de una opción europea.

Así, la función de densidad transitoria queda como sigue

$$f_{X_t|x_t}(x|X_t) = \frac{1}{\theta_2 x \sqrt{2\pi}(T - t)} \exp\left\{-\frac{1}{2}\left(\frac{\ln\left(\frac{x}{X_t}\right) - \left(\theta_1 - \frac{1}{2}\theta_2^2\right)(T - t)}{\theta_2\sqrt{T - t}}\right)^2\right\}.$$

Donde el valor esperado del precio de activo subyacente al vencimiento T , es decir, la esperanza del valor X_T condicionado al valor x_t , vendría dado por

$$\mathbb{E}[X_T|x_t] = X_t e^{\theta_1(T-t)}.$$

La varianza condicional sería

$$\text{Var}[X_T|x_t] = X_t^2 e^{2\theta_1(T-t)} e^{\theta_2^2(T-t)-1}.$$

En general, la densidad transitoria quedaría

$$p_\theta(t, y|x_0) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log y - \mu)}{2\sigma^2}\right\},$$

donde la media condicional y la varianza condicional, respectivamente, serían

$$\mu = \log(x_0) + \left(\theta_1 - \frac{1}{2}\theta_2^2\right)t,$$

y

$$\sigma^2 = \theta_2^2 t.$$

Sólo resta añadir que un movimiento geométrico Browniano pertenece a la clase de procesos de difusión Markovianos pues los movimientos actuales de los activos son independientes de los movimientos del pasado.

Asimismo, comentar que pertenecen a esa familia de procesos de difusión denominados de Pearson como se verá más adelante.

2.3. Modelo de Cox-Ingersoll-Ross

Este modelo surge como una mejora del modelo de Vasicek, aquél en el que $\theta_1 = 0$ en el modelo de Ornstein-Uhlenbeck, quedando así

$$dX_t = -\theta_2 X_t dt + \theta_3 X_t dW_t,$$

puede producir tasas de interés negativas con probabilidad positiva. Con la publicación de «A Theory of the Term Structure of Interest Rates» (1985),² el modelo de Cox-Ingersoll-Ross soluciona este «problema» haciendo que las tasas de interés sean estrictamente mayores que cero con probabilidad uno.

El proceso de Cox-Ingersoll-Ross es la solución a la ecuación diferencial estocástica que sigue

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 \sqrt{X_t} dW_t, \quad (2.8)$$

con condición inicial $X_0 = x_0 > 0$; siendo $\theta_1, \theta_2, \theta_3 \in \mathbb{R}_+$. Este proceso es no-negativo, esto es, se podría alcanzar el estado cero; sin embargo, si $2\theta_1 > \theta_3^2$, el proceso es estrictamente positivo. Observamos que el parámetro drift es lineal e idéntico al del proceso de Ornstein-Uhlenbeck; por ello, estamos ante un proceso que igualmente presenta reversión a la media como en el de Ornstein-Uhlenbeck; también es perceptible que, si hacemos que $\theta_1 = 0$, estaríamos ante el proceso de Vasicek.

En esta comparación con el modelo, se desprende que es en la varianza donde estriba la principal diferencia.

Así, supongamos que podemos modelar la dinámica del proceso a través de una ecuación diferencial estocástica más genérica, del tipo

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 X_t^{\theta_4} dW_t,$$

con la misma condición inicial y con $\theta_1, \theta_2, \theta_3 \in \mathbb{R}_+$.

Donde se aprecia que, si $\theta_4 = 0$, tendríamos el modelo de Ornstein-Uhlenbeck, con el parámetro difusión siendo una constante positiva, θ_3 . Pero, si $\theta_4 = \frac{1}{2}$, nos encontraríamos ante el modelo de Cox-Ingersoll-Ross, con un parámetro de difusión variable igual a $\theta_3 \sqrt{X_t}$, cuyo valor dependerá del espacio de estados, es decir, a medida que X_t aumenta, la desviación estándar aumenta; efectuando un analogismo con un proceso que modela la dinámica de los tipos de interés, la varianza se incrementa proporcionalmente a $\theta_3^2 X_t$ por unidad de tiempo, esto es, la varianza se incrementa a medida que las tasas de interés aumentan en una proporción de θ_3^2 ; en consecuencia, existe una mayor aleatoriedad en este modelo. Sin embargo, en

²John C. Cox, Jonathan E. Ingersoll y Stephen A. Ross, *Econometrica* Vol. 53, Núm 2 (Marzo, 1985), pág. 385-407, publicada por The Econometric Society.

el ambiente económico-financiero, esta varianza no captura efectivamente el llamado efecto leverage, no obstante, este no es el objeto de nuestro estudio.

Vamos a determinar, a continuación, la distribución condicional del proceso. Ya éramos conocedores de que si $2\theta_1 > \theta_3^2$, el proceso es estrictamente positivo, pero, además, el proceso es estacionario y, adicionalmente, la densidad de transición existe y se puede obtener explícitamente.

Para ello, consideremos el proceso de Vasicek ³

$$dX_t = -\theta_2 X_t dt + \theta_3 dW_t,$$

con $\theta_2, \theta_3 > 0$ y $\{W_t, t > 0\}$, un movimiento Browniano definido sobre un espacio fijo de probabilidad $(\Omega, \mathfrak{F}, \mathbb{P})$.

La solución explícita a este proceso es ⁴

$$X_t = x_0 e^{-\theta_2 t} + \theta_3 \int_0^t e^{-\theta_2(t-u)} dW_u. \quad (2.9)$$

La densidad de la distribución de X_t dado $X_0 = x_0$ posee como media y varianza condicional

$$\mathbb{E}_\theta(X_t | X_0 = x_0) = x_0 e^{-\theta_2 t}, \quad (2.10)$$

$$\text{Var}_\theta(X_t | X_0 = x_0) = \frac{\theta_3^2 (1 - e^{-2\theta_2 t})}{2\theta_2}, \quad (2.11)$$

donde se observa que, tanto el proceso de Vasicek como el de Ornstein-Uhlenbeck, obviamente poseen la misma media condicional pero diferente varianza condicional.

Supongamos que ahora tenemos n movimientos Brownianos independientes y definidos en el mismo espacio fijo de probabilidad $(\Omega, \mathfrak{F}, \mathbb{P})$, que denotaremos W_{1t}, \dots, W_{nt} . Supongamos n procesos de Vasicek, definidos de la forma

$$dX_{kt} = -\frac{1}{2}\theta_2 X_{kt} dt + \frac{1}{2}\theta_3 dW_{kt}, \quad (2.12)$$

³Aquél que es estrictamente una particularización del de Ornstein-Uhlenbeck, cuando el parámetro $\theta_1 = 0$, concretamente.

⁴Basta con tomar la solución explícita al proceso de Ornstein-Uhlenbeck 2.3 y sustituir el parámetro θ_1 por cero.

con $k = 1, 2, \dots, n$. Así, de las ecuaciones 2.11 y 2.9, se obtiene que

$$X_{kt} = e^{-\frac{\theta_2 t}{2}} \left(X_{k0} + \frac{1}{2} \theta_3 \int_0^t e^{\frac{\theta_2 s}{2}} dW_{ks} \right),$$

con $k = 1, 2, \dots, n$. Además, resulta que

$$\mathbb{E}_\theta(X_{kt} | X_{k0} = x_{k0}) = x_{k0} e^{-\frac{\theta_2 t}{2}} \equiv \mu_{k,t}$$

y que

$$Var_\theta(X_{kt} | X_{k0} = x_{k0}) = \frac{\theta_3^2 (1 - e^{-\theta_2 t})}{4\theta_2} \equiv \nu_t^2. \quad (2.13)$$

Consideremos el proceso

$$r_t = \sum_{k=1}^n X_{kt}^2, \quad (2.14)$$

entonces,

$$\frac{r_t}{\nu_t^2} \sim \chi^2(df, \delta_t), \quad (2.15)$$

donde $\delta_t = \frac{1}{\nu_t^2} \sum_{k=1}^n \mu_{k,t}^2$, es el parámetro de no centralidad.

Y $\frac{r_t}{\nu_t^2} \sim \chi^2(df, \delta_t)$, con df , los grados de libertad y con δ_t , el parámetro de no centralidad y df , los grados de libertad que determinaremos más adelante.

De 2.14 y 2.12, se sigue que

$$dr_t = \left(\frac{df\theta_3^2}{4} - \theta_2 r_t \right) dt + \theta_3 \sqrt{r_t} \sum_{k=1}^n \frac{X_{kt} dW_{kt}}{\sqrt{r_t}},$$

y definimos

$$dW_t \equiv \sum_{k=1}^n \frac{X_{kt} dW_{kt}}{\sqrt{r_t}},$$

llegamos a

$$dr_t = \left(\frac{df\theta_3^2}{4} - \theta_2 r_t \right) dt + \theta_3 \sqrt{r_t} dW_t, \quad (2.16)$$

donde $\{W_t, t > 0\}$ es un movimiento Browniano definido en el espacio fijo de probabilidad $(\Omega, \mathfrak{F}, \mathbb{P})$.

Recordemos que la ecuación 2.8 se podría reparametrizar como

$$dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t,$$

entonces, si se considera esta ecuación junto con 2.16, se tiene que

$$k\theta = \frac{df}{4}\theta_3^2,$$

así, por tanto,

$$df = \frac{4k\theta}{\theta_3^2}.$$

Además, el parámetro de no centralidad también se puede expresar como

$$\delta_t = \frac{1}{\nu_t^2}r_0e^{-kt}.$$

Por ende, si se considera 2.3, entonces se llega a

$$\frac{4k}{\sigma^2(1 - e^{-kt})} \cdot r_t \sim \chi^2\left(\frac{4k\theta}{\sigma^2}, \frac{4kr_0e^{-kt}}{\sigma^2(1 - e^{-kt})}\right).$$

Así, el proceso estocástico

$$Y_t = 2cX_t,$$

donde $c = \frac{2k}{\sigma^2(1 - e^{-kt})}$, tiene como distribución condicional $Y_t|Y_0 = y_0$, una chi cuadrado con grados de libertad df y δ_t como parámetro de no centralidad. Así pues, la densidad condicional de $X_t|X_0 = x_0$ se obtiene de la distribución $Y_t|Y_0 = y_0$, definida por

$$m(t, x_0) = \frac{k\theta}{k} + \left(x_0 - \frac{k\theta}{k}e^{-kt}\right),$$

y

$$v(t, x_0) = x_0 \frac{\sigma^2(e^{kt} - e^{-2kt})}{k} + \frac{k\theta\sigma^2(1 - e^{-2kt})}{2k^2}.$$

Donde $m(t, x_0)$ y $v(t, x_0)$ son la media y varianza condicional de la densidad transitoria, que van a ser de una enorme utilidad en el capítulo de las simulaciones o aproximaciones y, por ello, en la aplicabilidad de las ecuaciones diferenciales estocásticas en el ámbito de las finanzas.

2.4. Otros Modelos Paramétricos

En esta sección, se opta por desarrollar las principales características de modelos que son soluciones a ecuaciones diferenciales estocásticas que poseen un coeficiente drift no lineal o un coeficiente difusión de estructura más compleja.

En primer lugar se tratarán modelos que guardan una especial relación con el modelo de Cox-Ingersoll-Ross. Dado que la sección del modelo CIR se ha finalizado con su densidad condicional, el siguiente modelo que se presenta es el proceso de Ahn-Gao que es solución a la ecuación diferencial estocástica

$$dX_t = X_t(\theta_1 - (\theta_3^3 - \theta_1\theta_2)X_t)dt + \theta_3 X_t^{\frac{3}{2}} dW_t.$$

Como se observa es un proceso con coeficiente drift no lineal y un coeficiente difusión muy similar al del proceso CIR. El proceso de Ahn-Gao posee una densidad de transición explícita que es

$$p_\theta(t, y|x) = \frac{1}{y^2} p_\theta^{CIR}\left(t, \frac{1}{y} \middle| \frac{1}{x_0}\right),$$

donde $p_\theta^{CIR} = p_\theta(t, y|x_0) = ce^{-(u+v)} \left(\frac{u}{v}\right)^{\frac{q}{2}} I_q(2\sqrt{uv})$ es la densidad de transición del CIR, siendo $c = \frac{2\theta_2}{\sigma^2(1-e^{-\theta_2 t})}$, $u = cx_0 e^{-\theta_2 t}$, $v = cy$, $q = \frac{2\theta_1}{\theta_3^2} - 1$ e $I_q(\cdot)$ una función de Bessel modificada de primera clase de orden q .

Otros procesos estocásticos con coeficiente de difusión muy parecido al del CIR, aunque con coeficiente drift lineal son todos aquellos que pertenecen a la familia de modelos de Chan-Karolyi-Longstaff-Sanders que son soluciones a ecuaciones diferenciales estocásticas del tipo

$$dX_t = (\theta_1 + \theta_2 X_t)dt + \theta_3 X_t^{\theta_4} dW_t,$$

donde se puede percibir claramente que si $\theta_4 = \frac{1}{2}$ estaríamos ante un modelo CIR con

parámetro $\theta_2 = -\theta_2$ con lo que ya se podrían aplicar todas las propiedades intrínsecas al proceso CIR que ya se ha tratado. Básicamente, $\theta_1, \theta_3 \in \mathbb{R}_+$ y si $2\theta_1 > \theta_3^2$, el proceso es estrictamente positivo y estacionario con lo que tendría una densidad incondicional explícita con distribución chi cuadrado no centrada, y una densidad condicional con una varianza condicional idéntica a la del proceso CIR y una media condicional

$$m(t, x_0) = -\frac{\theta_1}{\theta_2} + \left(x_0 + \frac{\theta_1}{\theta_2} e^{\theta_2 t}\right).$$

Sin embargo, si $\theta_1 = 0$, $\theta_2 = 0$ y $\theta_3 = 1$, el proceso solución a la ecuación siguiente

$$dX_t = \theta_3 X_t dW_t,$$

es conocido por el modelo de Dothan.

Y si $\theta_4 = 1$, el modelo solución a la EDE $dX_t = (\theta_1 + \theta_2 X_t)dt + \theta_3 X_t dW_t$ es el proceso de Brennan-Schwartz.

Y si $\theta_1 = 0$ y $\theta_3 = 1$, estaríamos ante el modelo del movimiento geométrico Browniano, ya desarrollado.

Entre los modelos que presentan similitud con el proceso CIR se encuentran los procesos de difusión de Pearson que se caracterizan por poseer una densidad incondicional que pertenece al sistema de Pearson cuya función de densidad (de Pearson, de ahí el nombre del sistema), $p(x)$, es solución a una ecuación diferencial de la forma

$$\frac{p_x(x)}{p(x)} + \frac{a + x - \lambda}{b_2(x - \lambda)^2 + b_1(x - \lambda) + b_0},$$

donde b_2 , b_1 y b_0 tienen unas fórmulas explícitas que no detallaremos⁵; $p_x(x)$ es la derivada parcial de la función de Pearson respecto de x y λ es un parámetro de posición.

La densidad de Pearson adopta la forma $p(x) \propto \exp\left\{-\int \frac{x-a}{f(x)} dx\right\}$, donde la parte interesante está radicada en la función cuadrática $f(x) = b_2 x^2 + b_1 x + b_0$, que dependiendo de si tiene un discriminante positivo o no, las distribuciones pertenecientes al sistema de Pearson podrán tomar valores positivos o negativos, podrán estar acotadas, podrán ser simétricas o asimétricas, podrán tener colas ligeras o pesadas. La que realmente se usa en el ámbito financiero es la distribución de Pearson de tipo IV que, sin entrar en detalles, involucra a funciones

⁵Véase Wikipedia.

gamma y beta y que es parecida a una t de Student.

Estos procesos de difusión son soluciones a una EDE de la forma

$$dX_t = \theta(\mu - X_t)dt + \sqrt{2\theta(aX_t^2 + bX_t + c)}dW_t,$$

que si el discriminante y θ son positivos, el coeficiente de difusión estaría bien definido.

Como se puede apreciar, estos procesos con coeficiente drift lineal revierten a la media μ a velocidad θ poseen un coeficiente de difusión que es un polinomio de orden dos dependiente de los estados del proceso.

Asimismo, vemos que si el polinomio $(aX_t^2 + bX_t + c) = 1$, entonces el coeficiente de difusión $\sigma^2(x) = 2\theta$ y el proceso resultante es un proceso de difusión de Pearson pero con nombre propio, el modelo de Ornstein-Uhlenbeck, ya tratado.

Si $0 < \mu \leq 1$ y el coeficiente de difusión $\sigma^2(x) = 2\theta x$, resulta otro proceso de difusión de Pearson llamado modelo de Cox-Ingersoll-Ross, ya tratado también. En fin, bajo otras restricciones, se obtendrían otros procesos como el proceso de difusión de Jacobi, etc.

Seguidamente, se van a explicar muy brevemente otros tres modelos cuyo coeficiente de difusión es la unidad. Así, el proceso denominado hiperbólico es solución de la ecuación diferencial estocástica

$$dX_t = -\frac{\theta X_t}{\sqrt{1 + X_t^2}}dt + dW_t,$$

que es altamente aplicable en el contexto económico-financiero en el modelado de activos de renta variable.

Otro proceso con $\sigma^2 = 1$, es el radial de Ornstein-Uhlenbeck que es solución a una EDE del tipo

$$dX_t = (\theta X_t^{-1} - X_t)dt + dW_t,$$

con $\theta > 0$.

Su densidad de transición es conocida explícitamente que contiene una función de tipo sinusoidal que incluye una función de Bessel $I_\nu(\cdot)$, de orden ν .

Por último, aunque todos estos modelos anteriormente desarrollados no representan ni mucho menos la globalidad de los existentes, simplemente mencionar un proceso con forma

bimodal, que si se aplica el esquema de discretización de Milstein puede tener cierta aplicabilidad en las finanzas. Es solución de la ecuación diferencial estocástica $dX_t = (X_t - X_t^3)dt + dW_t$, con coeficiente drift no lineal.

Capítulo 3

Métodos de Simulación de Ecuaciones Diferenciales Estocásticas

Los métodos de simulación generalmente se basan en aproximaciones discretas de las soluciones continuas de las ecuaciones diferenciales estocásticas. Hasta ahora, hemos visto diferentes procesos paramétricos como soluciones explícitas a ecuaciones diferenciales estocásticas de variadas formas, en esta sección, se estudiarán los diversos métodos de aproximación más comúnmente empleados. En la simulación de soluciones, se abordarán desde los métodos más básicos hasta los más actuales.

Antes de continuar con nuestro estudio, estimamos conveniente presentar la transformación de Lamperti que es de especial interés tanto para los métodos de estimación como para los métodos de simulación. Básicamente, se trata de transformar una ecuación diferencial estocástica en otra con parámetro de difusión unitario. Así, supongamos una ecuación diferencial estocástica de la forma

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

donde se observa que el parámetro de difusión únicamente depende del espacio de estados, $\sigma(X_t)$.

Si aplicamos la transformación

$$F(X_t) = Y_t = \int_x^{X_t} \frac{1}{\sigma(u)} du,$$

donde x es cualquier estado del espacio de estados X .

La nueva ecuación diferencial estocástica sería

$$dY_t = b_Y(t, Y_t)dt + dW_t,$$

cuya solución es el proceso Y_t . Para obtener este uno, primero obtenemos las derivadas parciales de orden uno y derivada parcial de orden dos respecto; así, tenemos que

$$Y_t = f(t, x) = \int_x^{X_t} \frac{1}{\sigma(u)} du,$$

siendo sus derivadas parciales respecto de t y x

$f_t(t, x) = 0$, $f_x(t, x) = \frac{1}{\sigma(x)}$ y $f_{xx} = -\frac{\sigma_x(x)}{\sigma^2(x)}$, y al usar la fórmula de Itô con $f(t, x)$, se tiene que

$$\begin{aligned} df(t, x) &= 0 \cdot dt + f_x(t, X_t)dX_t + \frac{1}{2}f_{xx}(t, X_t)(dX_t)^2 \\ &= \frac{b(t, X_t)dt + \sigma(X_t)dW_t}{\sigma(X_t)} - \frac{1}{2} \frac{\sigma_x(X_t)}{\sigma^2(X_t)} (b(t, X_t)dt + \sigma(X_t)dW_t)^2 \\ &= \frac{b(t, X_t)}{\sigma(X_t)} dt + dW_t - \frac{1}{2} \sigma_x(X_t) \sigma^2(X_t) \sigma^2(x) dt. \end{aligned}$$

Por ende, se obtiene que

$$df(t, x) = dY_t = \left(\frac{b(t, X_t)}{\sigma(X_t)} - \frac{1}{2} \sigma_x(X_t) \right) dt + dW_t.$$

3.1. Métodos de Discretización de las Ecuaciones Diferenciales Estocásticas

Todos los métodos que se van a desarrollar están basados en la versión discretizada de una ecuación diferencial estocástica. En primer lugar, se presentará uno de los métodos más ampliamente usados, el esquema o método de Euler. Seguidamente, se introducirá el esquema de Milstein que añade un término de orden dos, para posteriormente analizar qué relación guardan entre ellos. En tercer lugar, se expone un método, el método de Predictor-Corrector, que considera a los parámetros como dependientes, no sólo del espacio de estados, sino del tiempo. Finalmente, se mostrará el segundo esquema de Milstein que requiere derivadas parciales primeras y segundas de los parámetros drift y difusión.

3.1.1. Esquema de Euler

Dado un proceso de Itô $\{X_t, 0 \leq t \leq T\}$ que es solución a la ecuación diferencial estocástica que sigue

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad (3.1)$$

con condición inicial no aleatoria $X_{t_0} = x_0$.

Sea la discretización de $[0, T]$, $\Pi_N = \Pi_N([0, T])$, igual a $0 = t_0 < t_1 < \dots < t_N = T$.

La aproximación de Euler de X es un proceso estocástico Y , continuo que satisface

$$Y(t_{i+1}) = Y(t_i) + b(t_i, Y(t_i))(t_{i+1}, t_i) + \sigma(t_i, Y(t_i))(W_{i+1}, W_i), \quad (3.2)$$

para $i = 0, 1, \dots, N-2, N-1$; con $Y(t_0) = X_{t_0} = X_0$.

Sea $\Delta t = t_{i+1} - t_i = \frac{1}{N}$, constante. Entre dos puntos de tiempo cualesquiera, el proceso se define de la forma siguiente

$$Y(t) = Y_{t_i} + \frac{t - t_i}{t_{i+1} - t_i}(Y(t_{i+1}) - Y(t_i)), \quad (3.3)$$

para cualquier $t \in [t_i, t_{i+1})$.

Para simular el proceso Y que, recordemos, es una aproximación de Euler de X que es solución de una ecuación diferencial estocástica de la forma de 3.1, bastaría con simular el incremento del proceso de Wiener (W_{i+1}, W_i) , pues los valores de $b(t, x)$ y de $\sigma(t, x)$ dependerían del proceso solución de la ecuación diferencial estocástica.

Así, sea un proceso de Ornstein-Uhlenbeck solución de 2.2. En este caso, $b(t, x) = (\theta_1 - \theta_2 x)$ y $\sigma(t, x) = \theta_3$. La aproximación de Euler de X , solución de 2.2 es un Y que satisface 3.2, de tal manera que únicamente tendríamos que simular el incremento del movimiento Browniano (W_{i+1}, W_i) , puesto que Y_{t_i} , $(\theta_1 - \theta_2 x)$ y θ_3 , o bien ya están fijados como Y_{t_i} , o bien inferidos como $b(t, x)$ y $\sigma(t, x)$. El incremento del proceso de Wiener se podría simular de la siguiente manera.

Se dividiría el intervalo $[0, T]$ en una malla de forma que $0 = t_0 < t_1 < \dots < t_N = T$ con $\Delta t = t_{i+1} - t_i$. Se fija $i = 1$ y $W(0) = W(t_1) = 0$ e itera el algoritmo siguiente¹:

¹Ver código de R, sección del movimiento Browniano.

32 Capítulo 3. Métodos de Simulación de Ecuaciones Diferenciales Estocásticas

Paso 1.- Se genera un nuevo número aleatorio z mediante una función $rnorm()$, pues $z \sim \mathbb{N}(0, 1)$. Paso 2.- Se hace $i = i + 1$. Paso 3.- Se fija $W(t_i) = W(t_{i-1} + z \cdot \sqrt{\Delta t})$. Paso 4.- Si $i \leq N$, volver al paso 1.

```
### --- Euler-Proceso OU --- ###
set.seed(1234)
theta <- c(4, 2, 2)
x0 <- 1
N <- 100
T <- 1 # Longitud del intervalo [0, T]
Delta <- T/N
Y <- numeric(N+1)
Y[1] <- x0
Z <- rnorm(N)
for(i in 1:N)
  Y[i+1] <- Y[i] + (theta[1]-theta[2]*Y[i])*Delta + theta[3]*sqrt(Delta)*Z[i]
Y <- ts(Y, start = 0, deltat=1/N)
plot(Y, main = "Aproximación Euler Proceso Ornstein-Uhlenbeck")
```

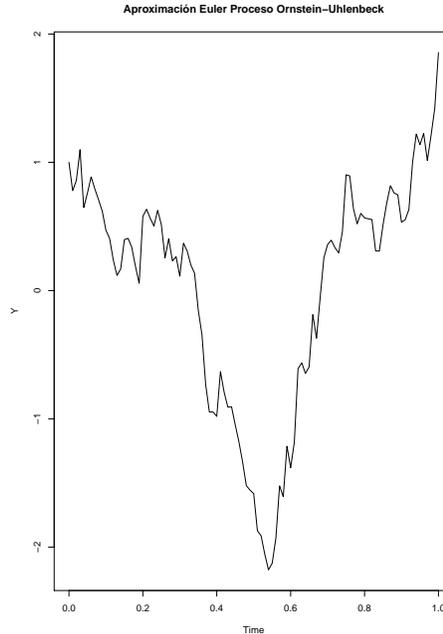


Figura 3.1: Trayectoria simulada del proceso de Ornstein-Uhlenbeck usando el esquema de Euler en lugar de la integral de Itô como ya se ha mostrado en la sección correspondiente. En el gráfico, se observa cómo el proceso alcanza valores positivos y negativos; característica ésta de particular trascendencia, pues, en la coyuntura actual de tipos de interés, está de rabiosa actualidad.

Y como ya se ha mencionado y se puede observar en el gráfico 3.1, entre dos puntos de tiempo cualesquiera, para cualquier $t \in [t_i, t_{i+1})$, la trayectoria se aproxima por interpolación lineal 3.3.

Si la solución de 2.2 es el proceso de Vasicek, lo único que variaría con respecto al anterior es $b(t, x) = -\theta_2 x$ y todo lo demás permanece exactamente igual.

```
### --- Euler-Proceso Vasicek --- ###
set.seed(1234)
theta <- c(4, 2, 2)
x0 <- 1
N <- 100
```

34 Capítulo 3. Métodos de Simulación de Ecuaciones Diferenciales Estocásticas

```
T <- 1 # Longitud del intervalo [0, T]
Delta <- T/N
Y <- numeric(N+1)
Y[1] <- x0
Z <- rnorm(N)
for(i in 1:N)
  Y[i+1] <- Y[i] + (-theta[2]*Y[i])*Delta + theta[3]*sqrt(Delta)*Z[i]
Y <- ts(Y, start = 0, deltat=1/N)
plot(Y, main = "Aproximación Euler Proceso Vasicek")
```

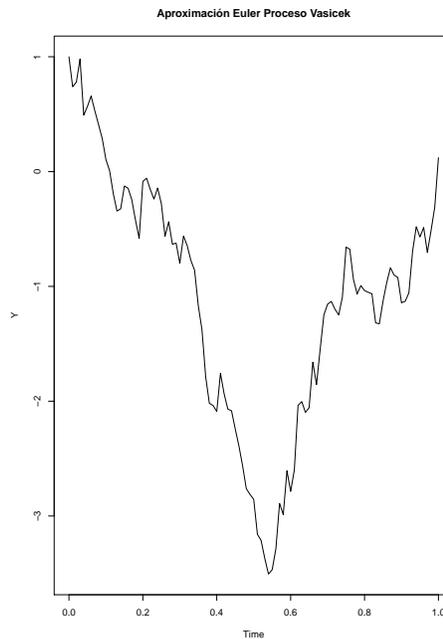


Figura 3.2: Trayectoria simulada del proceso de Vasicek que, como se aprecia, es muy parecida a la del modelo de Ornstein-Uhlenbeck; la diferencia surge de considerar al parámetro θ_1 de Ornstein-Uhlenbeck igual a cero.

Y si la solución de la ecuación 2.2 es el modelo de Black-Scholes-Merton, con $b(t, x) = \theta_1 x$ y $\sigma(t, x) = \theta_2 x$, la aproximación de Euler se efectuaría de forma análoga.

```
### --- Euler-Modelo de Black-Scholes-Merton --- ###
set.seed(1234)
theta <- c(4, 2, 2)
x0 <- 1
N <- 100
T <- 1 # Longitud del intervalo [0, T]
Delta <- T/N
Y <- numeric(N+1)
Y[1] <- x0
Z <- rnorm(N)
for(i in 1:N)
  Y[i+1] <- Y[i] + (theta[1]*Y[i])*Delta + theta[2]*Y[i]*sqrt(Delta)*Z[i]
Y <- ts(Y, start = 0, deltat=1/N)
plot(Y, main = "Aproximación Euler Modelo Black-Scholes-Merton")
```

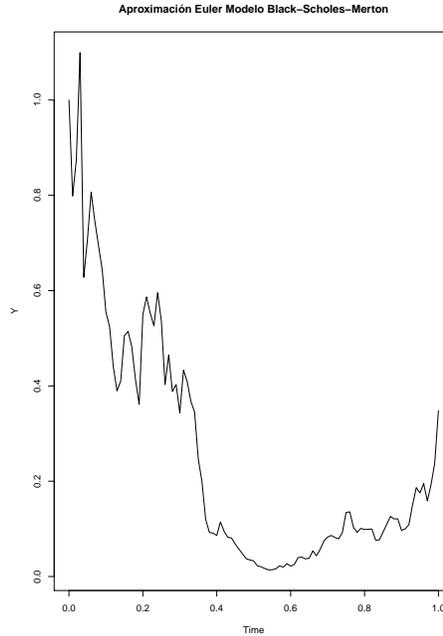


Figura 3.3: Trayectoria aproximada del modelo de Black-Scholes-Merton por la discretización de Euler.

En el caso de un proceso de Cox-Ingersoll-Ross solución de 2.2, con $b(t, x) = \theta_1 - \theta_2 x$ y $\sigma(t, x) = \theta_3 \sqrt{x}$, la simulación de Euler sólo afectaría a los incrementos del proceso de Wiener, como en los tres casos anteriores ya comentados.

```
### --- Euler-Proceso de Cox-Ingersoll-Ross --- ###
set.seed(1234)
theta <- c(4, 2, 2)
x0 <- 1
N <- 100
T <- 1 # Longitud del intervalo [0, T]
Delta <- T/N
Y <- numeric(N+1)
Y[1] <- x0
Z <- rnorm(N)
```

```

for(i in 1:N)
  Y[i+1] <- Y[i] + (theta[1]-theta[2]*Y[i])*Delta + theta[3]*sqrt(Y[i])*sqrt(Delta)*Z[i]
Y <- ts(Y, start = 0, deltat=1/N)
plot(Y, main = "Aproximación Euler Proceso Cox-Ingersoll-Ross")

```

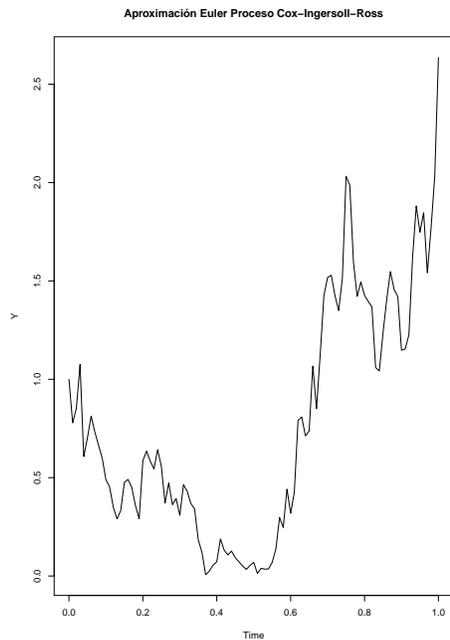


Figura 3.4: Trayectoria simulada del proceso CIR por esquema de aproximación de Euler. Destaca la presencia de valores estrictamente positivos como consecuencia de cumplirse $2\theta_1 > \theta_3^2$.

3.1.2. Esquema de Milstein: $\beta, \gamma = 1$

Este método simplemente incrementa la precisión del esquema de Euler usando el lema de Itô al añadir un término de orden dos. El esquema de Milstein utiliza la derivada parcial de orden dos del parámetro difusión $\sigma(t, x)$ respecto de x y sería como sigue

$$Y(t_{i+1}) = Y(t_i) + b(t_i, Y(t_i))(t_{i+1} - t_i) + \sigma(t_i, Y(t_i))(W_{i+1} - W_i) + \frac{1}{2}\sigma(t_i, Y(t_i))\sigma_x(t_i, Y(t_i))\{(W_{i+1} - W_i)^2 - (t_{i+1} - t_i)\},$$

Como se observa la única diferencia con el método de Euler es la entrada en escena del término $\frac{1}{2}\sigma(t_i, Y(t_i))\sigma_x(t_i, Y(t_i))\{(W_{i+1} - W_i)^2 - (t_{i+1} - t_i)\}$. Se puede deducir con facilidad que si la derivada de $\sigma(t, x)$ respecto de x es cero, estaríamos ante el esquema de Euler y, es únicamente en este caso donde Euler es de orden de convergencia fuerte $\gamma = 1$. Por ende, se recomienda usar el esquema de Milstein frente al de Euler en aquellos procesos soluciones a ecuaciones diferenciales estocásticas con el parámetro difusión dependiente del espacio de estados, pues, obviamente la derivada parcial respecto a x existiría y sería diferente de cero, y el orden de convergencia siempre sería superior al de Euler.

Se procederá a desarrollar algunos casos de procesos estocásticos donde el esquema de Milstein es más apropiado que el de Euler.

Así, se comenzará por el modelo del movimiento geométrico Browniano. Como ya se ha visto, éste es solución a la ecuación diferencial estocástica 2.7. En este caso, se tiene que $b(t, x) = \theta_1 \cdot x$, $\sigma(t, x) = \theta_2 \cdot x$ y $\sigma_x(t, x) = \theta_2$.

Entonces, usando el esquema de Milstein

$$Y(t_{i+1}) = Y(t_i) + \theta_1 \cdot Y(t_i)(t_{i+1} - t_i) + \theta_2 Y(t_i) \cdot (W_{i+1} - W_i) + \frac{1}{2}\theta_2 \cdot Y(t_i)\theta_2\{(W_{i+1} - W_i)^2 - (t_{i+1} - t_i)\}$$

Simplificando

$$\begin{aligned} Y(t_{i+1}) &= Y(t_i) + \theta_1 \cdot Y(t_i)\Delta t + \theta_2 \cdot Y(t_i)\Delta W_t + \frac{1}{2}\theta_2^2 \cdot Y(t_i)(\Delta W_t^2 - \Delta t) \\ &= Y(t_i)\left(1 + \left(\theta_1 - \frac{1}{2}\theta_2^2\right)\Delta t\right) + \theta_2 Y(t_i)\Delta W_t + \frac{1}{2}\theta_2^2 Y(t_i)(\Delta W_t)^2. \end{aligned}$$

Se sabe que los incrementos del proceso de Wiener, $\Delta W_t \sim \sqrt{\Delta t}Z$, donde $Z \sim \mathbb{N}(0, 1)$; por ello

$$Y(t_{i+1}) = Y(t_i) \left(1 + \left(\theta_1 - \frac{1}{2} \theta_2^2 \right) \Delta t \right) + \theta_2 Y(t_i) \sqrt{\Delta t} W_t Z + \frac{1}{2} \theta_2^2 Y(t_i) \Delta t Z^2$$

Simplificando, queda así

$$Y(t_{i+1}) = Y(t_i) \left(1 + \left(\theta_1 + \frac{1}{2} \theta_2^2 (Z^2 - 1) \right) \Delta t + \theta_2 \sqrt{\Delta t} Z \right)$$

El código de R que simula una trayectoria del modelo de Black-Scholes-Merton discretizada por el esquema de Milstein sería:

```
### --- Milstein-Modelo de Black-Scholes-Merton
set.seed(1234)
theta <- c(4, 2, 2)
x0 <- 1
N <- 100
T <- 1 # Longitud del intervalo [0, T]
Delta <- T/N
Y <- numeric(N+1)
Y[1] <- x0
Z <- rnorm(N)
for(i in 1:N)
  Y[i+1] <- Y[i]*(1+(theta[1]+1/2*theta[2]^2*(Z[i]^2-1)))*Delta + theta[2]*sqrt(Delta)*Z[i]
Y <- ts(Y, start = 0, deltat=1/N)
plot(Y, main = "Aproximación Milstein Modelo Black-Scholes-Merton")
```

O también

```
### Milstein B-S-M
require(sde)
theta1 = 0.1934661
theta2 = 0.1618493

drift <- expression(theta1 * x)
difussion <- expression(theta2*x)
```

```
part_deriv <- expression(theta2)

set.seed(234)
X <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
            N = 2547, method = "milstein")
plot(X, main = "Black-Scholes-Merton (MILSTEIN)")
```

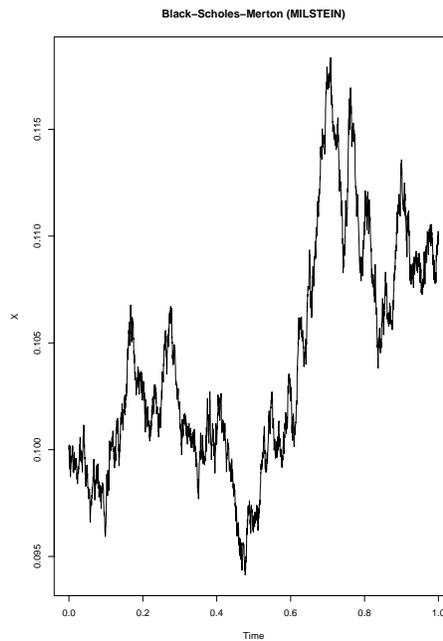


Figura 3.5: Trayectoria simulada del movimiento geométrico Browniano a través de la discretización de Milstein.

Si el proceso es el Cox-Ingersoll-Ross la solución a una ecuación diferencial estocástica de la forma 2.8, se tiene que

$$b(t, x) = (\theta_1 - \theta_2 \cdot x), \sigma(t, x) = \theta_3 \cdot \sqrt{x} \text{ y } \sigma_x(t, x) = \frac{\theta_3}{\sqrt{x}}.$$

La aproximación por el esquema de Milstein sería

$$Y(t_{i+1}) = Y(t_i) + (\theta_1 - \theta_2 \cdot Y(t_i))\Delta t + \theta_3 \cdot \sqrt{Y(t_i)}\Delta W_t + \frac{1}{2}\theta_3 \cdot \sqrt{Y(t_i)} \frac{\theta_3}{\sqrt{Y(t_i)}} \{(\Delta W_t)^2 - \Delta t\}$$

Se sabe que los incrementos del proceso de Wiener, $\Delta W_t \sim \sqrt{\Delta t}Z$, donde $Z \sim \mathbb{N}(0, 1)$ y simplificando, se obtiene

$$Y(t_{i+1}) = \left\{ \theta_1 + Y(t_i)(1 - \theta_2) + \frac{1}{2}\theta_3^2(z^2 - 1) \right\} \Delta t + \theta_3 \sqrt{Y(t_i)} \sqrt{\Delta t} Z$$

Para concluir con este análisis, consideramos necesario destacar que el resultado, al transformar procesos soluciones de ecuaciones diferenciales estocásticas y aplicar el esquema de aproximación de Euler sobre el proceso transformado, es el mismo que cuando se usa el método de Milstein en la discretización del proceso original; así, el orden de convergencia que se obtiene es igual a Δt . Veamos esto implementado en dos códigos de R,

```
par(mfrow=c(1,2))
### CIR (MILSTEIN)
require(sde)
theta1 = 0.1733261
theta2 = 0.1934661
theta3 = 0.1618493

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(234)
X <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
            N = 2547, method = "milstein")
plot(X, main = "Cox-Ingersoll-Ross (MILSTEIN)")

### CIR-3 (EULER sobre la transformada Y = sqrt(X_{t}))
```

42 Capítulo 3. Métodos de Simulación de Ecuaciones Diferenciales Estocásticas

```
set.seed(12345)
theta1 = 0.1733261
theta2 = 0.1934661
theta3 = 0.1618493

drift <- expression((theta1 - theta2 * x^2 - theta3^2/4)/(2*x))
difussion <- expression(theta3/2)
Y <- sde.sim(X0=sqrt(0.1), drift = drift, sigma = difussion, N = 2547)
plot(Y^2, main = "C-Ing.-Ross (EULER Transform.)")
```

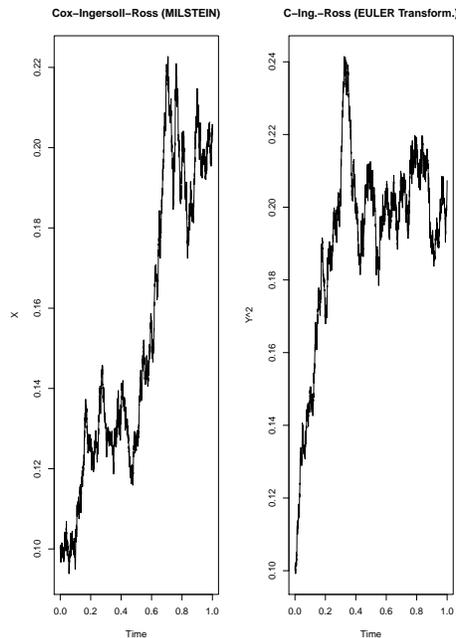


Figura 3.6: Trayectorias aparentemente diferentes que, en realidad, son idénticas pero a una escala diferente.

Como se puede observar en la figura 3.6, parece que las trayectorias que devuelven ambos esquemas son diferentes, sin embargo, son las mismas pero a distinta escala (i.e., $Y^2 = X$).

3.1.3. Método de Predictor-Corrector

Supongamos una ecuación diferencial estocástica genérica de la forma

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

con $\{W_t, t > 0\}$ un proceso de Wiener o movimiento Browniano.

Se percibe que tanto b como σ dependen del tiempo, t , y del estado del proceso, X_t . Sin embargo, los esquemas de simulación explicados en los dos apartados anteriores consideran que los coeficientes b y σ no varían con el tiempo.

Puesto que los coeficientes dependen de X_t , y se está simulando X_t , el método que se presenta a continuación, en una primera fase, fase llamada predictor, aproxima el espacio de estados del proceso, así

$$Y(\tilde{t}_{i+1}) = Y(t_i) + b(t_i, Y(t_i))\Delta t + \sigma(t_i, Y(t_i))\sqrt{\Delta t}Z,$$

donde $Z \sim \mathbb{N}(0, 1)$.

Hasta ahora, con respecto a un esquema de Euler, no hay diferencia, la única distinción es de naturaleza nominativa (predictor).

Seguidamente, se le añaden unos pesos, $0 \leq \alpha \leq 1$ y $0 \leq \eta \leq 1$ y se calcula el predictor del coeficiente b , esto es, \tilde{b}

$$\tilde{b}(t_i, Y(t_i)) = b(t_i, Y(t_i)) - \eta\sigma(t_i, Y(t_i))\sigma_x(t_i, Y(t_i)),$$

donde $\sigma_x(t_i, Y(t_i))$ es la derivada parcial de primer orden respecto de x .

El corrector, se calcula como sigue

$$Y(t_{i+1}) = Y(t_i) + \left(\alpha\tilde{b}(t_{i+1}, Y(\tilde{t}_{i+1})) + (1 - \alpha)\tilde{b}(t_i, Y(t_i)) \right) \Delta t \\ + \left(\eta\sigma(t_{i+1}, Y(\tilde{t}_{i+1})) + (1 - \eta)\sigma(t_i, Y(t_i)) \right) \sqrt{\Delta t}Z.$$

Se ve que si $\eta = \alpha = 0$, los términos $\eta\sigma(t_{i+1}, Y(\tilde{t}_{i+1}))$ y $\alpha\tilde{b}(t_{i+1}, Y(\tilde{t}_{i+1}))$ se anularían, quedando el esquema de Euler.

Grosso modo, se trata de un esquema de discretización de Euler al que se corrige con unos pesos.

3.1.4. Segundo Esquema de Milstein

Este método añade una serie de términos a la aproximación consistentes en las derivadas parciales respecto de x de primer y segundo orden de los coeficientes drift y difusión. Está pensado para procesos estocásticos soluciones a ecuaciones diferenciales estocásticas con coeficientes de drift y difusión no lineales. Es un esquema que Milstein propuso en su día y resulta ser como sigue

$$Y_{i+1} = Y_i + \left(b - \frac{1}{2}\sigma\sigma_x\right)\Delta t + \sigma Z\sqrt{\Delta t} + \frac{1}{2}\sigma\sigma_x\Delta t Z^2 + \Delta t^{\frac{3}{2}}\left(\frac{1}{2}b\sigma_x + \frac{1}{2}b_x\sigma + \frac{1}{4}\sigma^2\sigma_{xx}\right)Z + \Delta t^2\left(\frac{1}{2}bb_x + \frac{1}{4}b_{xx}\sigma^2\right). \quad (3.4)$$

Donde b_x, σ_x son las derivadas de orden uno de b y σ respecto de x ; donde b_{xx}, σ_{xx} son las derivadas de orden dos de b y σ respecto de x , y $Z \sim \mathbb{N}(0, 1)$.

```
### --- Segundo Esquema de Milstein
set.seed(112233)
theta <- c(0.1733261, 0.1934661, 0.1618493)
x0 <- 1
t0 <- 1
N <- 100
T <- 1 # Longitud del intervalo [0, T]
Delta <- T/N

d1 <- expression(theta[1] - theta[2] * x)
d1.x <- expression(-theta[2])
d1.xx <- 0
s1 <- expression(theta[3]*sqrt(x))
s1.x <- expression(1/sqrt(x))
s1.xx <- expression(-1/(2*sqrt(x^3)))
```

```

sde.sim.milstein2 <- function(X0, t0, Delta, N, d1, d1.x, d1.xx, s1, s1.x, s1.xx){
  X <- numeric(N+1)
  Y <- numeric(N+1)
  sD <- sqrt(Delta)
  Z <- rnorm(N, sd=sD)
  X[1] <- X0
  Y[1] <- X0
  for(i in 2:(N+1)){
    X[i] <- X[i-1]+d1(t[i-1], X[i-1])*Delta+s1(t[i-1], X[i-1])*z[i-1]+
      0.5*s1(t[i-1], X[i-1])*s1.x(t[i-1], X[i-1])*(Z[i-1]^2-Delta) +
      Delta^(3/2)*(0.5*d1(t[i-1], X[i-1])*s1.x(t[i-1], X[i-1]))+
      0.5*d1.x(t[i-1], X[i-1])*s1(t[i-1], X[i-1])+0.25*s1(t[i-1], X[i-1])^2*
      s1.xx(t[i-1], X[i-1])*Z[i-1]+Delta^2*(0.5*d1(t[i-1], X[i-1])*d1.x(t[i-1], X[i-1])+
      0.25*d1.xx(t[i-1], X[i-1])*s1(t[i-1], X[i-1])^2)

  }
  return(X)
}
X <- sde.sim.milstein2(X0=x0, t0, Delta, N, d1, d1.x, d1.xx, s1, s1.x, s1.xx)

```

En el código anterior, se muestra cómo se puede aplicar la discretización del segundo esquema de Milstein. Vemos la absolutamente necesaria aportación de las derivadas de orden uno y dos de los coeficientes drift y de difusión, así como el segundo esquema de Milstein 3.4.

3.2. Métodos de Aproximación Local

También llamados como métodos de linealización local y consisten básicamente en aproximar el coeficiente drift con una función lineal. Son esquema válidos tanto para ecuaciones diferenciales estocásticas homogéneas o no. El primer método que se presenta posee una elevada utilidad para procesos estocásticos soluciones de ecuaciones cuyo coeficiente de difusión es constante. El siguiente esquema abre un poco más el abanico de usos, siendo interesante para aquellos procesos que son soluciones de ecuaciones con coeficientes de difusión que dependen del espacio de estados.

3.2.1. Método Ozaki

Se trata de un esquema de aproximación que, en apariencia, carece de recorrido en términos de aplicabilidad, pues únicamente se podría usar en ecuaciones diferenciales estocásticas homogéneas; sin embargo, nada más lejos de la realidad, ya que mediante la transformación de Lamperti, podemos transformar una ecuación diferencial estocástica con una volatilidad dependiente de los estados, en una volatilidad unitaria (este proceso de transformación aunque aparentemente simplista, dota de una practicidad al profesional que, a falta de todo, algo es bastante). Supongamos la ecuación diferencial estocástica de la forma

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

con $\{W_t, t > 0\}$ un proceso de Wiener.

Se asume que la derivada parcial de primer orden respecto de x del coeficiente drift, $b_x(x)$ es constante localmente, en el intervalo $[t, t + \Delta t]$; tras varios cálculos, se llega a la ecuación diferencial

$$x_{t+\Delta t} = x_t + \frac{b(x_t)}{b_x(x_t)}(e^{b_x(x_t)\Delta t} - 1).$$

Si se asume que $b(x)$ se aproxima por $K_t x$, con K_t , depende de t a través de X_t , es constante en $[t, t + \Delta t]$; entonces la solución a la ecuación sería

$$X_t = X_t e^{K_x \Delta t} + \sigma \int_t^{t+\Delta t} e^{K_x(t+\Delta t-u)} dW_u,$$

con

$$K_x = \frac{1}{\Delta t} \log \left(1 + \frac{b(x)}{x b_x(x)} (e^{b_x(x)\Delta t} - 1) \right). \quad (3.5)$$

La densidad condicional de $X_{t+\Delta t}$ dada X_t sigue una distribución normal con media condicional igual a $x + \frac{b(x)}{x b_x(x)}(e^{b_x(x)\Delta t} - 1)$ y varianza condicional dependiente de K_x e igual a $\sigma^2 \frac{e^{2K_x \Delta t} - 1}{2K_x}$.

Se puede apreciar que la media condicional existiría aunque el coeficiente drift fuese constante, no así la varianza condicional, pues K_x , el logaritmo de la expresión 3.5 sería cero y, por ello, tanto la constante K_x como la varianza condicional no estarían bien definidas.

Finalmente, supongamos un proceso de Vasicek puro de la forma

$$dX_t = -\theta_2 X_t dt + \theta_3 dW_t.$$

Con $\theta_2 X_t = -3x$ y $\theta_3 = 2$.

En este supuesto, con un coeficiente drift lineal como es en el caso de un proceso de Vasicek, el método de Euler y el método Ozaki son iguales; así, mediante la implementación del código siguiente se observa la superposición de ambas trayectorias.

```
require(sde)
set.seed(123456)
X <- sde.sim(drift=expression(-3*x), sigma=expression(2), method="euler") # Método de Euler
plot(X, lty=1, col="green", main= "Drift Lineal: Euler = Ozaki")
set.seed(123456)
Y <- sde.sim(drift=expression(-3*x), sigma=expression(2), method="ozaki") # Método de Ozaki
lines(as.numeric(time(Y)), Y, col="red", lty=3)
```

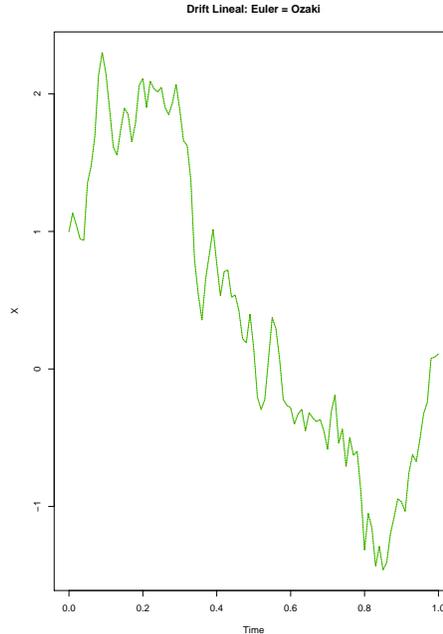


Figura 3.7: Las aproximaciones mediante el método de Euler y el de Ozaki de un proceso de difusión como el modelo de Vasicek con coeficiente drift lineal, arrojan trayectorias idénticas.

3.2.2. Método Shoji-Ozaki

Este método de linealización local amplía considerablemente el objeto de su actuación. Se ha visto en la sección anterior que el método Ozaki se usa en ecuaciones diferenciales estocásticas homogéneas con parámetro de difusión constante y con coeficiente drift dependiente únicamente del espacio de estados. Sin embargo, el método de Shoji-Ozaki puede ser empleado en la discretización de una ecuación diferencial estocástica del tipo

$$dX_t = b(t, X_t)dt + \sigma dW_t \quad (3.6)$$

la cual, como se puede observar, se trata de una ecuación diferencial estocástica no homogénea que posee un coeficiente drift que depende de t y de X , y un parámetro de difusión, σ , constante. Se asume que $b(t, X_t)$ es dos veces diferenciable en x y en t . Este método estudia el comportamiento de $b(t, X_t)$ localmente, esto es, en $[s, s + \Delta s]$ donde Δ es muy pequeña.

No vamos a entrar en los detalles del proceso de obtención del resultado de la discretización mediante Shoji-Ozaki que es como sigue²

$$X_{s+\Delta s} = B(X_s)X_s + C(X_s)Z, \quad (3.7)$$

donde

$$B(X_s) = 1 + \frac{b(e^{b_x \Delta s} - 1)}{X_s b_x} + \frac{\frac{\sigma^2}{2}(b_{xx} + b_t)}{X_s b_x^2} (e^{b_x \Delta s} - 1 - b_x \Delta s), \quad (3.8)$$

y

$$C(X_s) = \sigma \sqrt{\frac{e^{2b_x \Delta s} - 1}{2b_x}}, \quad (3.9)$$

siendo b la notación de $b(s, X_s)$, $b_x = b_x(s, X_s)$, $b_{xx} = b_{xx}(x, X_s)$ y Z es una normal estándar de media cero y varianza uno.

Así,

$$X_{s+\Delta s} | X_s = x \sim N(B(x)x, C^2(x)). \quad (3.10)$$

Vemos que la varianza condicional depende del estado previo, igual que sucede con el método Ozaki, no así con el de Euler.

En el caso de una ecuación diferencial estocástica homogénea, los métodos de Euler, Ozaki y Shoji-Ozaki coinciden, siendo este último más estable con una Δ grande. Veámoslo en el siguiente código, suponiendo que nuestra EDE sea

```
bx <- expression((5-11*x+6*x^2-x^3))
x0 <- 5
Dt <- 0.1
par(mfrow=c(2,3))
set.seed(12345)
X <- sde.sim(drift=bx, delta = Dt, X0=x0, method="euler")
plot(X, main="Esquema de Euler")
```

²«Estimation for nonlinear Stochastic Differential Equations by a Local Linearization Method», Shoji, L. y Ozaki, T. (1998), Stochastic Anal. Applications, 16, pág. 733-752.

```
set.seed(12345)
Y <- sde.sim(drift=bx, delta = Dt, X0=x0, method="ozaki")
plot(Y, main="Método de Ozaki")
set.seed(12345)
Z <- sde.sim(drift=bx, delta=Dt, X0=x0, method="shoji")
plot(Z, main="Método de Shoji-Ozaki")
```

Tal y como el gráfico 3.8 muestra, una ecuación diferencial estocástica homogénea con parámetro de difusión constante, un coeficiente drift lineal y un Δ pequeño, discretizada por cualquiera de los tres métodos, Euler, Ozaki o Shoji-Ozaki, arroja trayectorias absolutamente idénticas.

Como ya sabemos de secciones precedentes, el esquema de Euler con un Δ grande, no funciona, tiende a explotar. Lo mismo ocurriría con Ozaki que, aunque ligeramente mejor pues la aproximación es mejor que la aproximación de una constante, también explota. Sin embargo, el método de linealización local de Shoji-Ozaki se muestra sólido y estable. Veámoslo con el código que sigue con una $\Delta = 0.50$

```
bx <- expression((5-11*x+6*x^2-x^3))
x0 <- 5
Dt <- 0.5
par(mfrow=c(2,3))
set.seed(12345)
X <- sde.sim(drift=bx, delta = Dt, X0=x0, method="euler")
plot(X, main="Esquema de Euler")
set.seed(12345)
Y <- sde.sim(drift=bx, delta = Dt, X0=x0, method="ozaki")
plot(Y, main="Método de Ozaki")
set.seed(12345)
Z <- sde.sim(drift=bx, delta=Dt, X0=x0, method="shoji")
plot(Z, main="Método de Shoji-Ozaki")
```

Si nos encontramos ante un EDE no-homogénea con un coeficiente de difusión que dependa

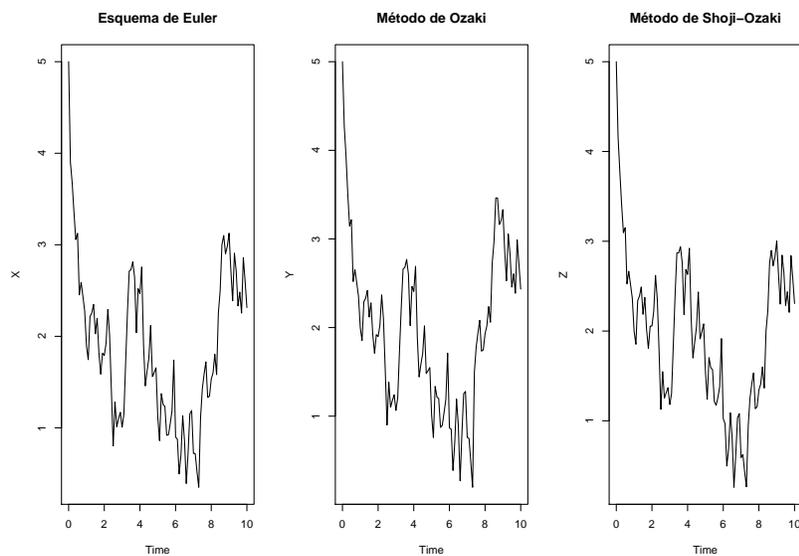


Figura 3.8: Trayectorias idénticas de una EDE homogénea con parámetro de difusión unidad (consecuencia de una transformación de Lamperti), coeficiente drift lineal y Δ pequeño, obtenidas por los métodos de Euler, Ozaki y Shoji-Ozaki.

del espacio de estados (un proceso CIR, un movimiento geométrico Browniano, por ejemplo)

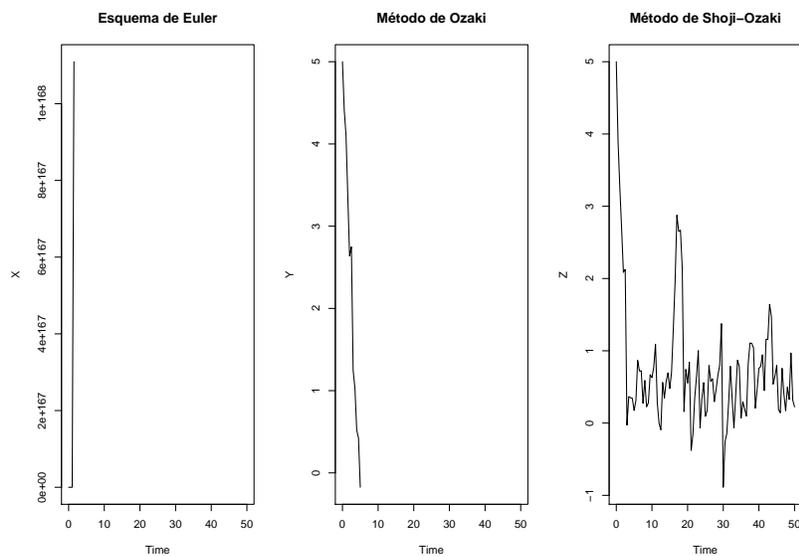


Figura 3.9: Con una $\Delta = 0.5$, los métodos de Euler y de Ozaki explotan. Sin embargo, el método de Shoji-Ozaki se muestra estable con deltas grandes (continuamos en las mismas circunstancias, esto es, una EDE homogénea con parámetro de difusión unidad -consecuencia de una transformación de Lamperti- y coeficiente drift lineal).

del tipo

$$dX_t = b(t, X_t)dt + \sigma(X_t)dW_t, \quad (3.11)$$

se asume que $b(t, X_t)$ es dos veces diferenciable en x y en t ; en este caso, se aplicaría la transformación de Lamperti, esto es, $\sigma(X_t) = 1$, de modo que el proceso ya es discretizable.

3.3. Métodos de Estimación Paramétrica

En la presente sección se van a desarrollar los diversos métodos de inferencia paramétrica de los valores teóricos de procesos de difusión con muestras en tiempos discretos. Se va a suponer que el tiempo entre dos valores consecutivos de la muestra es fijo y se denotará por Δ y también se asume que el intervalo $[0, n\Delta = T]$, se incrementa a medida que n aumenta. Asimismo, se va requerir que el proceso sea estacionario y/o ergódico. Además, se supone que el proceso se observa en tiempo discreto; así, $t_i = i\Delta_i, i = 0, 1, \dots, n$, y $T = n\Delta_n$; la tasa de muestreo será, o bien, constante, $\Delta_i = \Delta$, o bien, $\max_i \Delta_i < \Delta$ para un delta fijo.

Por un lado, si estuviésemos en un caso continuo, parece bastante sencillo calcular eficientemente los parámetros. Así, los parámetros que conforman la parte de difusión del proceso pueden ser calculados

$$\langle X, X \rangle_t = \int_0^t \sigma^2(X_s, \theta) ds,$$

para todo $t \geq 0$.

El resto de los parámetros de la parte drift pueden estimarse por el método de máxima verosimilitud. Una vez que el coeficiente de difusión se conoce, la función de verosimilitud de X es la derivada de Radon-Nykodým que se expresa más abajo. Previamente, se necesita comprender el teorema de Girsanov que simplemente es un teorema de cambio de medida en procesos estocásticos. La versión que se va a exponer es la siguiente

Sean las ecuaciones diferenciales estocásticas

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

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

$$dX_t = \sigma(X_t)dW_t,$$

con condiciones de inicio X_0 aleatorias y cuyas funciones de densidad son $f_1(\cdot)$, $f_2(\cdot)$ y $f_3(\cdot)$, con el mismo dominio (también se podría asumir que las condiciones iniciales sean deterministas y, entonces, tendría que asumirse que fuesen iguales a una constante); se tiene que $0 \leq t \leq T$. Se denota por P_1 , P_2 y P_3 , las medidas de probabilidad inducidas por las soluciones de las ecuaciones diferenciales estocásticas.

Se asume ³ que la condición global de Lipschitz y la condición de crecimiento lineal se satisfacen, así como cualquier otra condición necesaria para garantizar la existencia y unicidad de la solución de cada una de las ecuaciones diferenciales estocásticas anteriores. Entonces las tres medidas de probabilidad, P_1 , P_2 y P_3 , son equivalentes y sus respectivas derivadas de Radon-Nikodým son

$$\frac{dP_1}{dP_3}(X) = \frac{f_1(X_0)}{f_3(X_0)} \exp \left\{ \int_0^T \frac{b_1(X_s)}{\sigma^2(X_s)} dX_s - \frac{1}{2} \int_0^T \frac{(b_1(X_s))^2}{\sigma^2(X_s)} ds \right\}$$

y

$$\frac{dP_2}{dP_1}(X) = \frac{f_2(X_0)}{f_1(X_0)} \exp \left\{ \int_0^T \frac{b_2(X_s) - b_1(X_s)}{\sigma^2(X_s)} dX_s - \frac{1}{2} \int_0^T \frac{(b_2(X_s))^2 - (b_1(X_s))^2}{\sigma^2(X_s)} ds \right\}$$

Se puede ver que $b_3(X_s) = 0$.

Por ende, la derivada de Radon-Nikodým, esto es, la función de verosimilitud de X , denotada por $L_T(\theta)$, es

$$L_T(\theta) = \exp \left\{ \int_0^T \frac{b(X_s, \theta)}{\sigma^2(X_s)} dX_s - \frac{1}{2} \int_0^T \frac{(b(X_s, \theta))^2}{\sigma^2(X_s)} ds \right\},$$

θ se estimaría al maximizar $L_T(\theta)$.

No obstante, no se nos debe olvidar que no estamos en un caso de tiempo continuo, por tanto, la función de verosimilitud de X en tiempo discreto y condicionada a X_0 es

$$L_n(\theta) = \prod_{i=1}^n p_\theta(\Delta, X_i | X_{i-1}) p_\theta(X_0),$$

³Ver «Statistics of Random Processes», volumen 1, de Lipster, R.S., Shiriyayev, A.N. (1997), Edit. Springer-Verlag.

donde $p_\theta(X_0) = 1$, pues el peso relativo de ésta en la verosimilitud disminuye a medida que n aumenta. Si se desea la log-verosimilitud, entonces se tendría que

$$\log L_n(\theta) \propto \prod_{i=1}^n \log p_\theta(\Delta, X_i | X_{i-1}).$$

Así que, una vez expuestos algunos teoremas y conceptos básicos para una buena comprensión, en esta sección, primero, se va a proceder a desarrollar los métodos de inferencia de verosimilitud exacta para aquellos modelos paramétricos donde la densidad de transición se conoce. Seguidamente, se trata con otro modo de obtener estimadores en el caso de modelos paramétricos con densidad de transición desconocida, mediante los métodos de pseudo-verosimilitud.

3.3.1. Métodos de Estimación de Verosimilitud Exacta

Ya se han desarrollado algunos procesos estocásticos que son solución a ecuaciones diferenciales estocásticas. Estos modelos, como también ya se ha visto, poseen una densidad transitoria o condicional conocida. Entre los más usados, destacan el proceso de Ornstein-Uhlenbeck, de Vasicek, de Black-Scholes-Merton y de Cox-Ingersoll-Ross, fundamentalmente. En este apartado, se estudiará el método de inferencia de verosimilitud exacta. Se necesitan una serie de condiciones suficientes para que estos estimadores sean consistentes, asintóticamente Gaussianos, eficientes e invariantes. Algunas de ellas ya se han expuesto en apartados anteriores: la condición global de Lipschitz 1.16 y la condición de crecimiento lineal 1.17. Existen algunas otras asunciones de cumplimiento suficiente como

$$\inf_x \sigma^2(x, \theta) > 0$$

$$\sup_t \mathbb{E}|X_t|^k < \infty,$$

para $k > 0$

La condición por la que el coeficiente de difusión sea estrictamente positivo siempre se pide en todos los procesos estocásticos analizados en este documento. La condición de la existencia de todos los momentos de orden k y, a la que se debe adicionar otra acerca de los coeficientes de drift y de difusión y sus derivadas (de hasta orden tres) en θ sean suaves en x .

Veamos algunos códigos de R para el cálculo directo de los estimadores o para estimación mediante optimización numérica. En secciones anteriores ya se ha tratado con suficiencia, el proceso de Vasicek como solución a la ecuación diferencial estocástica

$$dX_t = -\theta_2 X_t dt + \theta_3 dW_t,$$

con $X_0 = x_0$; con $\theta_2 \in R$ y $\theta_3 \in \mathbb{R}_+$.

Además, si $\theta_2 > 0$, el proceso presenta ergodicidad. En apartados anteriores, también se ha presentado su densidad condicional con su media condicional 2.10 y su varianza condicional 2.11. Con toda esta información, se está en disposición de confeccionar el siguiente código para el cálculo explícito de los estimadores de máxima verosimilitud de θ_2 y θ_3 , y compararlo con los obtenidos numéricamente mediante la función de R, *mle()*.

```
### --- VASISCEK Estim. Explícitos de MV y Estimad. numéricos de MV
require(stats4)
require(sde)
set.seed(2345)

# Densidad Condicional
dcOU <- function(x, t, x0, theta, log = FALSE){
  m.cond <- x0*exp(-theta[2]*t)
  v.cond <- theta[3]^2*(1-exp(-2*theta[2]*t))/(2*theta[2])
  dnorm(x, mean=m.cond, sd=sqrt(v.cond), log=log)
}

# Función de Verosimilitud
OU.lik <-function(theta1, theta2, theta3){
  n <- length(X)
  dt <- deltat(X) -sum(dcOU(X[2:n], dt, X[1:(n-1)], c(theta1, theta2, theta3 ), log=TRUE))
}

# Simulación de una Solución de Ornstein-Uhlenbeck con theta1 = 0
X <- sde.sim(model = "OU", theta = c(0, 3, 2), N=1000, delta = 1)
# Estimación por Máxima Verosimilitud de los parámetros
```

```

fit <- mle(OU.lik, start = list(theta2=1.5, theta3=1), fixed = list(theta1=0), method = "L-B
fit
# Coefficients:
#   theta1   theta2   theta3
# 0.000000 4.287168 2.358410
prof <- profile(fit)
par(mfrow=c(1,3))
plot(prof) # Perfiles de Verosimilitud para los Parámetros del Modelo Vasicek
vcov(fit) # Matriz Hessiana Invertida: la Varianza Estimada de los Estimadores
confint(fit) # Intervalos de Confianza de los Estimadores

# Cálculo Explícito de los Estimadores de Máxima Verosimilitud
n <- length(X)
tmp.sum <- sum(X[1:(n-1)]*X[2:n])
dt <- deltat(X)
theta2.hat <- ifelse(tmp.sum>0, -log(tmp.sum/sum(X[1:(n-1)]^2))/dt, NA) # Estimador MV de th
theta3.hat <- sqrt(2*theta2.hat/((n-1)*(1-exp(-2*dt*theta2.hat)))*sum((X[2:n]-X[1:(n-1)]*exp
theta2.hat
# [1] 4.286993
theta3.hat
# [1] 2.358362

```

Los estimadores de θ_2 y θ_3 son consistentes; los resultados arrojados por ambos métodos son muy similares.

Como se puede observar en el código anterior, el cálculo exacto de los estimadores de máxima verosimilitud, $\hat{\theta}_2$ y $\hat{\theta}_3$, se pueden obtener explícitamente mediante

$$\hat{\theta}_{2,n} = -\frac{1}{\Delta} \log \left[\frac{\sum_{i=1}^n X_{i-1} X_i}{\sum_{i=1}^n (X_{i-1})^2} \right], \quad (3.12)$$

para una tasa de muestreo Δ fija.

3.12 está definida si y sólo si $\sum_{i=1}^n X_{i-1} X_i > 0$.

$$\hat{\theta}_{3,n} = \frac{2\hat{\theta}_{2,n}}{n(1 - e^{-2\hat{\theta}_{2,n}})} \sum_{i=1}^n (X_1 - X_{i-1}e^{-\Delta\hat{\theta}_{2,n}})^2.$$

Si en lugar de un proceso de Vasicek puro, estamos ante un proceso de Ornstein-Uhlenbeck solución a la ecuación diferencial estocástica 2.2, con la condición inicial determinista, $X_0 = x_0$, con $\theta_1, \theta_2 \in \mathbb{R}$ y $\theta_3 \in \mathbb{R}_+$.

Si $\theta_2 > 0$, el proceso oscila alrededor de un valor y es ergódico. Su densidad de transición exacta es una Gaussiana con la media condicional 2.4 y la varianza condicional 2.5 que ya se han visto en la sección correspondiente. Así, análogamente al proceso de Vasicek, con esta información se procede a confeccionar el código siguiente

```
### --- Estim. Numéricos del Modelo de Ornstein-Uhlenbeck
require(stats4)
require(sde)
set.seed(2345)

# Densidad Condicional
theta = c(3, 1, 2)
dcOU <- function(x, t, x0, theta, log = FALSE){
  m.cond <- theta[1]/theta[2]+(x0 - theta[1]/theta[2])*exp(-theta[2]*t)
  v.cond <- theta[3]^2*(1-exp(-2*theta[2]*t))/(2*theta[2])
  dnorm(x, mean=m.cond, sd=sqrt(v.cond), log=log)
}

# Función de Verosimilitud
OU.lik <-function(theta1, theta2, theta3){
  n <- length(X)
  dt <- deltat(X)
  -sum(dcOU(X[2:n], dt, X[1:(n-1)], c(theta1, theta2, theta3 ), log=TRUE))
}

# Simulación de una Solución de Ornstein-Uhlenbeck
X <- sde.sim(model = "OU", theta = c(3, 1, 2), N=1000, delta = 1)
# Estimación por Máxima Verosimilitud de los parámetros
fit <- mle(OU.lik, start = list(theta1=1, theta2=0.5, theta3=1),
```

```
method = "L-BFGS-B", lower = c(-Inf, 0, 0))
summary(fit)
fit
# Coefficients:
#      Estimate Std. Error
# theta1 3.173695 0.26605606
# theta2 1.087535 0.08824101
# theta3 2.032978 0.07486703

# -2 log L: 3359.217

prof <- profile(fit)
par(mfrow=c(1,3))
plot(prof) # Perfiles de verosimilitud para los Parámetros del Modelo Ornstein-Uhlenbeck
vcov(fit) # Matriz Hessiana Invertida: la varianza estimada de los Estimadores
#      theta1      theta2      theta3
# theta1 0.07078583 0.022712984 0.015311427
# theta2 0.02271298 0.007786476 0.005249063
# theta3 0.01531143 0.005249063 0.005605072
confint(fit) # Intervalos de Confianza de los Estimadores
#      2.5 %    97.5 %
# theta1 2.6906517 3.744709
# theta2 0.9279102 1.277673
# theta3 1.8972347 2.193111
```

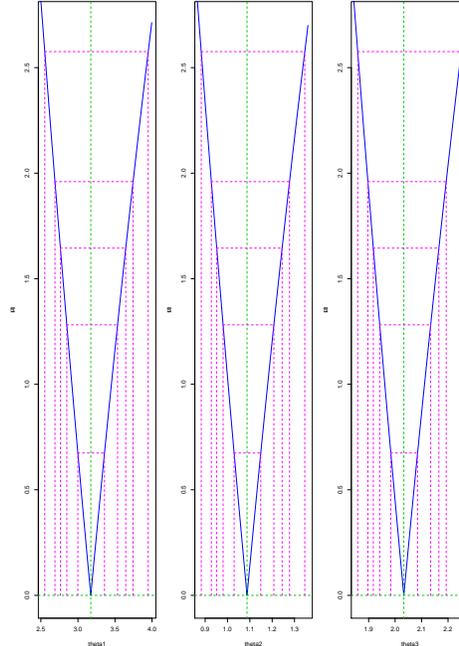


Figura 3.10: Perfiles de Verosimilitud de los estimadores de un modelo de Ornstein-Uhlenbeck.

Para el caso de un movimiento geométrico Browniano solución a una ecuación de la forma 2.7, se tendría, como ya se ha desarrollado en el apartado correspondiente, una densidad de transición log-normal cuya media y desviación estándar del logaritmo de variable son

$$ml(t, x) = \log x + \left(\theta_1 - \frac{1}{2}\theta_2^2\right)t$$

y

$$sl(t, x) = \sqrt{t}\theta_2.$$

Así pues, se puede elaborar el código siguiente:

```
### --- MGB Estim. Estimad. Numéricos de MV
require(stats4)
require(sde)
set.seed(2345)
```

```

# Densidad Condicional
dcMGB <- function(x, t, x0, theta, log = TRUE){
  m.cond <- log(x0) + (theta[1]-theta[2]^2/2)*t
  v.cond <- theta[2]*sqrt(t)
  dlnorm(x, meanlog = m.cond, sdlog = sqrt(v.cond), log=TRUE)
  if(!log)
    lik <- exp(lik)
  lik
}
# Función de Verosimilitud
MGB.lik <- function(theta1, theta2){
  n <- length(X)
  dt <- deltat(X) -sum(dcMGB(X[2:n], t=dt, x0=X[1:(n-1)]), theta=c(theta1, theta2), log=TRUE)
}
# Simulación de una Solución de Black-Scholes-Merton
X <- sde.sim(model = "BS", theta = c(0.5,0.2), N=1000, delta = 0.01)
# Estimación por Máxima Verosimilitud de los parámetros
fit <- mle(MGB.lik, start = list(theta1=1, theta2=1),
method = "L-BFGS-B", lower = c(0.01,0.01))

```

Por último, el proceso de Cox-Ingersoll-Ross solución de 2.8 tiene una densidad de transición que sigue una distribución χ^2 no centrada que puede ser expresada en términos de funciones de Bessel reescaladas exponencialmente. Así

```

# Función de Bessel Reescalada Exponencialmente
expBes <- function(x,nu){
  mu <- 4*nu^2
  A1 <- 1
  A2 <- A1*(mu- 1)/(1*(8*x))
  A3 <- A2*(mu- 9)/(2*(8*x))
  A4 <- A3*(mu-25)/(3*(8*x))
  A5 <- A4*(mu-49)/(4*(8*x))

```

62 Capítulo 3. Métodos de Simulación de Ecuaciones Diferenciales Estocásticas

```
A6 <- A5*(mu-81)/(5*(8*x))
A7 <- A6*(mu-121)/(6*(8*x))
(1/sqrt(2*pi*x))*(A1-A2+A3-A4+A5-A6+A7) }
```

La densidad de transición se puede escribir como sigue

$$p_{\theta}(t, y|x) = ce^{-u-v} \left(\frac{u}{v}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}),$$

recordemos que en el caso del modelo CIR, $x, y \in \mathbb{R}_+$.

La función de log-verosimilitud sería

$$l_i(\theta) = \log c - (u + v) + \frac{q}{2} \log \left(\frac{u}{v}\right) + \log I_q(2\sqrt{uv}).$$

Donde $c = \frac{2\theta_2}{\theta_3^2(1-e^{-\theta_2 t})}$, $u = cxe^{-\theta_2 t}$, $q = \frac{2\theta_1}{\theta_3^2} - 1$ y $v = cy$, figuran definidas en el siguiente código para la densidad condicional y la función de verosimilitud del proceso de Cox-Ingersoll-Ross

```
# Densidad Condicional del Modelo de Cox-Ingersoll-Ross
dcCIR <- function(x,t,x0,theta,log=FALSE){
  c <- 2*theta[2]/((1-exp(-theta[2]*t))*theta[3]^2)
  ncp <- 2*c*x0*exp(-theta[2]*t)
  df <- 4*theta[1]/theta[3]^2
  u <- c*x0*exp(-theta[2]*t)
  v <- c*x
  q <- 2*theta[1]/theta[3]^2-1
  lik <- log(c)-(u+v)+(q/2)*log(v/u)+log(expBes(2*sqrt(u*v),q))+2*sqrt(u*v)
  if (!log){lik <- exp(lik)}
  lik
}
```

y la función de verosimilitud del proceso de Cox-Ingersoll-Ross, así como la simulación de una solución del proceso CIR y la estimación numérica de los correspondientes parámetros de máxima verosimilitud

```
# Función de Verosimilitud
```

```
CIR.lik <- function(theta1,theta2,theta3){
  n <- length(X)
  dt <- deltat(X)
  -sum(dcCIR(x=X[2:n],t=dt,x0=X[1:(n-1)],theta=c(theta1,theta2,theta3),log=TRUE)) }

# Simulación de una solución del proceso CIR
X <- sde.sim(X0=0.1, model = "CIR", theta = c(0.2, 0.06, 0.15), N=2500, delta=0.1)
# Estimación de los parámetros de Máxima Verosimilitud
fit <- mle(CIR.lik, start = list(theta1=0.1, theta2=0.1, theta3=0.3),
method = "L-BFGS-B", lower=c(0.001,0.001,0.001), upper=c(1,1,1))
fit
# Coefficients:
#   theta1      theta2      theta3
# 0.21015798 0.06144055 0.15153237
prof <- profile(fit)
par(mfrow=c(1,3))
plot(prof) # Perfiles de Verosimilitud para los Parámetros del Modelo CIR
vcov(fit)
#           theta1      theta2      theta3
# theta1 1.349847e-03 4.175624e-04 3.837064e-06
# theta2 4.175624e-04 1.575815e-04 1.186823e-06
# theta3 3.837064e-06 1.186823e-06 4.600544e-06
confint(fit)
#           2.5 %      97.5 %
# theta1 0.13824116 0.28248951
# theta2 0.03686419 0.08612572
# theta3 0.14741809 0.15583090
```

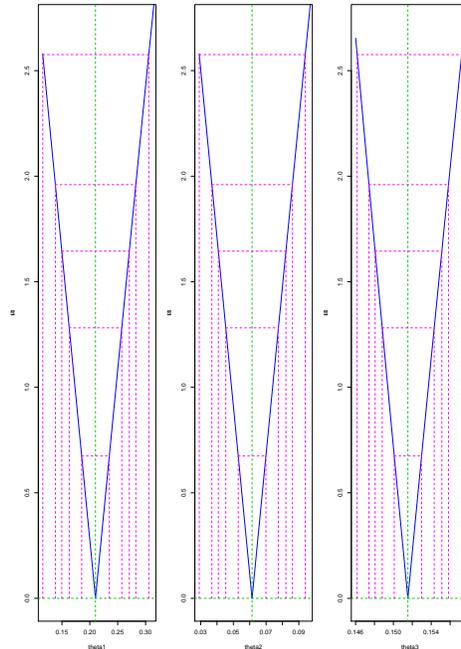


Figura 3.11: Perfiles de verosimilitud de los estimadores de un modelo de Cox-Ingersoll-Ross.

3.3.2. Métodos de Pseudo-Verosimilitud

En lugar de simular directamente la densidad condicional, estos métodos dan un pequeño rodeo. Primero, simulan la trayectoria del proceso para, posteriormente, aproximar la verosimilitud de la versión discretizada del mismo. Esta discretización se puede efectuar, como ya se ha visto en la sección correspondiente, por diversos esquemas, cuyo aplicabilidad dependerá de la homogeneidad o no de la ecuación diferencial estocástica. Así, en lo que sigue se abordarán los métodos de aproximación más conocidos que están muy relacionado con los conceptos ya revisados de los esquemas de Euler, Milstein, Shoji y Ozaki-Shoji.

Estos métodos siguen básicamente el siguiente proceso: se tiene, como siempre, un proceso que es solución de una ecuación diferencial estocástica. Se procedería seguidamente a su discretización por el esquema más adecuado. Posteriormente, se obtiene, cuando exista, la densidad de transición del proceso discretizado, para, a continuación, estimar por máxima verosimilitud los parámetros.

Se empezará con el esquema de discretización de Euler que, recordemos, funciona apropia-

damente siempre que Δt sea muy pequeña, de otro modo, aparecerá sesgo en la estimación; además, se asume que σ es estrictamente positiva y constante, no depende del espacio de estados ni del tiempo. Así,

- Proceso solución a la ecuación diferencial estocástica $dX_t = b(X_t, \theta) + \sigma(X_t, \theta)dW_t$, con $\sigma(X_t, \theta) = \sigma$.
- Por el esquema de discretización de Euler, $X_{t+\Delta t} - X_t = b(X_t, \theta)(t, t + \Delta t) + \sigma(X_t, \theta)(W_{t+\Delta t} - W_t)$ ⁴.
- Densidad de transición del proceso discretizado, $p_\theta(t, y|x) \sim \mathbb{N}(b(x, \theta), \sigma(x, \theta))$.
- La log-verosimilitud del proceso discretizado es equivalente a $\sum_{i=1}^n (X_i - X_{i-1}) - b(X_t, \theta) - \frac{\Delta}{2} \sum_{i=1}^n (b(X_t, \theta))^2$.

Este es el proceso básico para entender el método de pseudo-verosimilitud por Euler. Así, si Δt es muy pequeña, un proceso de Ornstein-Uhlenbeck solución a una ecuación diferencial estocástica de la forma

$$dX_t = b(X_t, \theta) + \sigma dW_t$$

se podría discretizar por Euler, para luego obtener su densidad condicional y estimar por máxima verosimilitud los parámetros. Ya conocemos su densidad condicional definida por su media condicional 2.4 y su varianza condicional 2.5 y, asimismo, sabemos que maximizar la log-verosimilitud del proceso discretizado es equivalente a maximizar la función 3.3.2.

En el supuesto de que $\sigma(X_t, \theta)$ no fuese constante y dependiese de un espacio de estados, i.e. el movimiento geométrico Browniano, se podría aplicar el método de pseudo-verosimilitud de Euler siempre que, o bien se efectúe la transformación de Lamperti haciendo que σ alcance el valor unitario, o bien se necesitaría un estimador del coeficiente de difusión para estimar por máxima verosimilitud los parámetros del coeficiente drift. Un estimador consistente de $\sigma(X_t, \theta)$ sería

$$\hat{\sigma}^2 = \frac{1}{n\Delta} \sum_{i=1}^n (X_i - X_{i-1})^2.$$

Si se usa un esquema de aproximación como el de Milstein, que incorpora un término de segundo orden al esquema de Euler, se estaría ante el método Elerian. La existencia de una forma cerrada de la densidad condicional para el esquema de Milstein fue debido a O. Elerian⁵, según el cual, dada la existencia de una única solución a una ecuación diferencial

⁴Se asume que $b(X_t, \theta)$ y $\sigma(X_t, \theta)$ permanecen constantes en $[t, t + \Delta t]$.

⁵Ver Teorema 2.1 de «A note of the existence of a closed form conditional density for the Milstein scheme»,

estocástica con coeficientes drift y de difusión tanto dependientes del tiempo como del espacio de estados, y dadas algunas condiciones adicionales, que no se van a tratar aquí, se llega a que la densidad condicional de un esquema de Milstein no es Gaussiana ni simétrica y es como sigue

$$p_{\theta}(t, y|x) = \frac{z^{-\frac{1}{2}} \cosh(\sqrt{Cz})}{|A|\sqrt{2\pi}} e^{-\frac{C+z}{2}},$$

donde

$$A = \frac{\sigma(x)\sigma_x(x)t}{2}$$

$$B = -\frac{\sigma(x)}{2\sigma_x(x)} + x + b(x)t - A$$

$$C = \frac{1}{(\sigma_x)^2(x)t}$$

$$z = \frac{y - B}{A}$$

donde

$\sigma_x(x)$ es la derivada parcial de $\sigma(t, x)$ respecto de x . Para $\sigma_x(x) \neq 0$ y $z > 0$.

El código a implementar para la densidad de transición de un esquema de Milstein sería

```
### --- Densidad Condicional de un Esquema de Milstein y Verosimilitud de Elerian
dcElerian(x, t, x0, t0, theta, d, s, sx, log = FALSE){
  A <- s(t0, x0, theta)*sx(t0, x0, theta)/2
  B <- -s(t0, x0, theta)/(2*sx(t0, x0, theta))+x0+d(t0, x0, theta)*t-A
  C <- 1/(sx(t0, x0, theta)^2)*t
  z <- (x - B)/A
  lik <- (exp(-(C+z)/2)*(exp(sqrt(C*z))+exp(-sqrt(C*z)))/(2*sqrt(z)*abs(A)*sqrt(2*pi)))
  if(log)
    lik <- log(lik)
  lik
}
```

Supongamos un proceso de Cox-Ingersoll-Ross solución de la EDE

$$dx_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 \sqrt{X_t} dW_t,$$

donde $\theta_1, \theta_2 \in \mathbb{R}$ y $\theta_3 \in \mathbb{R}_+$. Si $\theta_2 > 0$, el proceso es estacionario y ergódico.

En este caso, para poder mostrar cómo se comportan los diferentes métodos de pseudo-verosimilitud que hemos expuesto anteriormente, se debe proceder al cálculo de $b_x(t, x)$, $\sigma_x(x)$, $b_{xx}(t, x)$ y $\sigma_{xx}(x)$, las derivadas parciales de primer y segundo orden de los coeficientes de drift y difusión respecto de x , y también, el cálculo de $b_t(t, x)$, la derivada de primer orden respecto de t .

Supongamos que se desea estimar el parámetro θ_2 de una trayectoria simulada del modelo CIR con parámetros $(\theta_1, \theta_2, \theta_3) = (0.5, 0.2, \sqrt{0.05})$. Así, si aplicamos el método de pseudo-verosimilitud de Euler y el método de Elerian, se implementaría el código siguiente:

```
### --- Método de Pseudo-Verosimilitud de Euler y Elerian
b <- function(t,x,theta) theta[1]-theta[2]*x
bx <- function(t,x,theta) -theta[2]
bxx <- function(t,x,theta) 0
bt <- function(t,x,theta) 0
s <- function(t,x,theta) theta[3]*sqrt(x)
sx <- function(t,x,theta) theta[3]/(2*sqrt(x))
sxx <- function(t,x,theta) -theta[3]/(4*x^(3/2))

dcEULER <- function(x, t, x0, theta, log = FALSE){
  m.condEu <- x0*(1-theta[2]*t) + theta[1]*t
  v.condEu <- t*theta[3]^2
  dnorm(x, mean = m.condEu, sd = sqrt(v.condEu), log = log)
}

EULER.lik <- function(theta){
  sum(dcEULER(X[2:n], t[2:n], X[1:(n-1)], t[1:(n-1)],
c(0.5,theta,sqrt(0.05)), b, s, TRUE), na.rm = TRUE)
}
```

```
b2 <- function(t,x,theta){
  (4*theta[1]-theta[3]^2)/(2*x*theta[3]^2)-theta[2]*x/2
}
b2x <- function(t,x,theta){
  -theta[2]/2-(4*theta[1]-theta[3]^2)/(2*x^2*theta[3]^2)
}
b2xx <- function(t,x,theta){
  (4*theta[1]-theta[3]^2)/(x^3*theta[3]^2)
}
b2t <- function(t,x,theta) 0
s2 <- function(t,x,theta) 1
s2x <- function(t,x,theta) 0
s2xx <- function(t,x,theta) 0

ELERIAN.lik <- function(theta){
  sum(dcEULER(X[2:n], t[2:n], X[1:(n-1)], W[1:(n-1)], t[1:(n-1)],
c(0.5,theta,sqrt(0.05)), b2, s2, s2x, TRUE), -0.5*log(X[2:n]*0.05), na.rm = TRUE)
}
# Verosimilitud Exacta
TRUE.lik <- function(theta){
  sum(dcCIR(X[2:n], deltat(X), X[1:(n-1)], c(0.5,theta,sqrt(0.05)), TRUE), na.rm = TRUE)
}

pTRUE <- function(x) TRUE.lik
pEULER <- function(x) EULER.lik
pELERIAN <- function(x) ELERIAN.lik

set.seed(12345)
X1 <- sde.sim(model = "CIR", theta=c(0.5, 0.2, sqrt(0.05)), X0=2, delta = 0.001, N=50000)
xx <- seq(0.001, 0.4, length=50)

par(mfrow=c(2,2))
```

```

est <- NULL
for(dt in c(4,2,1,0.5)){
  X <- window(X1, deltat=dt)
  W <- 2*sqrt(X)/sqrt(0.05)
  t <- as.numeric(time(X))
  n <- length(X)
  cat(sprintf("numero de observaciones: %d, Delat=%3.2f\n", n, dt))
  dTRUE <- sapply(xx, pTRUE)
  dEULER <- sapply(xx, pEULER)
  dELERIAN <- sapply(xx, pELERIAN)

  mx <- max(c(dTRUE, dEULER, dELERIAN, na.rm = TRUE))
  mn <- min(c(dTRUE, dEULER, dELERIAN, na.rm = TRUE))

  matplot(xx, cbind(dTRUE, dEULER, dELERIAN), type = "l", ylim = c(mn,mx), xlab="",
ylab="aproxim.", main = sprintf("N=%d, Delta=%3.2f", n, dt), lty = 1:3,
col = 1:3)
  legend(0.15*0.6*(mn+mx), lty = 1:3, col = 1:3, legend = c("TRUE","EULER","ELERIAN"))

  tmp <- c(n, dt,
           optimize(pTRUE, c(0.01,0.4))$max,
           optimize(pEULER, c(0.01,0.4))$max,
           optimize(pELERIAN, c(0.01,0.4))$max)
  est <- rbind(est, tmp)
}

dimnames(est)[[2]] <- c("N", "Delta", "TRUE", "EULER", "ELERIAN")
dimnames(est)[[1]] <- 1:4
print(est)
par(mfrow=c(1,1))

```

Tal y como ya se ha desarrollado en secciones anteriores, existen otros enfoques alterna-

tivos para aproximar la solución a una ecuación diferencial estocástica, como los métodos de linealización local de Ozaki y Shoji-Ozaki. Estos métodos proporcionan una opción adicional a la hora de obtener una densidad condicional tanto si se trata de una ecuación diferencial estocástica homogénea de la forma

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

(donde se usaría el método de Ozaki, pues σ es constante. También se podría aplicar el método de pseudo-verosimilitud de Euler).

Como si se trata de una ecuación diferencial estocástica de la forma

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

(donde se aplicaría el esquema de Shoji-Ozaki, pues, aunque σ es constante, el coeficiente drift, $b(t, X_t)$, a diferencia del de Ozaki, depende no sólo del espacio de estados, sino también del tiempo).

Por ende, todos aquellos procesos estocásticos soluciones a ecuaciones diferenciales estocásticas cuyo coeficiente de difusión no es constante, deben ser transformados por Lamperti para, así, poder usar alguno de estos dos métodos.

Así, análogamente a lo realizado con el método de pseudo-verosimilitud de Euler y de Elerian, supongamos un proceso de Cox-Ingersoll-Ross solución de la EDE

$$dx_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 \sqrt{X_t} dW_t,$$

donde $\theta_1, \theta_2 \in \mathbb{R}$ y $\theta_3 \in \mathbb{R}_+$. Si $\theta_2 > 0$, el proceso es estacionario y ergódico.

En este caso, para poder mostrar cómo se comportan los diferentes métodos de pseudo-verosimilitud que hemos expuesto anteriormente, se debe proceder al cálculo de $b_x(t, x)$, $\sigma_x(x)$, $b_{xx}(t, x)$ y $\sigma_{xx}(x)$, las derivadas parciales de primer y segundo orden de los coeficientes de drift y difusión respecto de x , y también, el cálculo de $b_t(t, x)$, la derivada de primer orden respecto de t .

3.3.3. Métodos de Aproximación Directa de la Verosimilitud

En lugar de métodos que simulan la trayectoria de un proceso estocástico previamente discretizado, se conocen otros que aproximan la verosimilitud directamente.

Entre los que se van a tratar en este apartado, destaca el esquema cuya principal idea surge de la aproximación de la densidad condicional del proceso estocástico mediante una expansión similar a la efectuada por Harald Cramér ⁶ quién asevera que la suma de n variables independientes e idénticamente distribuidas, bajo ciertas circunstancias, convergerán en distribución.

Este método se conoce con el nombre de la expansión de los polinomios de Hermite cuyo precursor fue Ait-Sahalia quién aproximó la densidad condicional de un proceso y trató las asíntotas en el caso i.i.d. cuando $\Delta \rightarrow 0, n \rightarrow +\infty$. Ait-Sahalia, en lugar de expandir directamente la densidad objetivo alrededor de una normal estándar como había hecho Cramér, transforma primero la densidad objetivo en algo lo más semejante posible a una distribución normal, para, posteriormente, expandir la densidad transformada mediante los polinomios de Hermite.

Para una implementación efectiva de este esquema, es conveniente considerar que las observaciones son discretas. Por tanto, sean X_0, \dots, X_i observaciones discretas de un proceso X .

Nuestro objetivo es la densidad de transición o condicional siguiente

$$p_X(\Delta, X_i | X_{i-1}, \theta) = \sigma(x, \theta)^{-1} p_Y^{(j)(\Delta, F(x) | F(x_0), \theta)}, \quad (3.13)$$

con $j = 1, \dots, 6$.

La ecuación 3.13 se podría reexpresar como

$$p_X(\Delta, X_i | X_{i-1}, \theta) = \frac{\Delta^{-\frac{1}{2}}}{\sigma(X_i, \theta)} \cdot p_Z^{(j)(\Delta, \Delta^{-\frac{1}{2}}(F(X_i) - F(X_{i-1})) | F(X_{i-1}), \theta)}, \quad (3.14)$$

donde $p_Z^{j(\Delta, z | y_0, \theta)} = \phi(z) \sum_{j=0}^J \nu_Z^{(j)(\Delta, z | y_0, \theta)} H_j(z)$ y donde $\nu_Z^{(j)(\Delta | y_0, \theta)}$ puede ser expresada en función de los momentos de la difusión. No entraremos en más detalles innecesarios, pues lo realmente importante de este esquema es conocer el parámetro drift del proceso transformado,

⁶«On Asymptotic Expansions for Sums of Independent Variables with a Limit Stable Distribution», Cramér, H., Sankhyā: The Indian Journal of Statistics, Series A (1961-2002), Vol. 25, No. 1 (Jul., 1963), pp. 13-24

que será denotado por μ_Y , sus derivadas hasta el orden 6. Comentar también que se precisa conocer la función F , la transformada de Lamperti ($Y = F(X)$) que tiene que ser creciente e invertible, pues

$$\mu_Y(y, \theta) = \frac{b(F^{-1}(y), \theta)}{\sigma(F^{-1}(y), \theta) - \frac{1}{2}\sigma_x(F^{-1}(y), \theta)}. \quad (3.15)$$

Como se observa en la última ecuación, μ_Y depende del parámetro drift y de difusión del proceso original X .

Seguidamente se muestra un escueto resumen de los parámetros esenciales de la teoría que se acaba de exponer para los procesos de difusión más sencillos. Así, tenemos:

Ornstein-Uhlenbeck:

$$F(x) = \frac{x}{\theta_3}$$

$$\mu_Y(y; \theta) = \frac{\theta_1\theta_2}{\theta_3} - \theta_1 y$$

Cox-Ingersoll-Ross:

$$F(x) = \frac{2\sqrt{x}}{\theta_2}$$

$$\mu_Y(y; \theta) = \frac{4\theta_1\theta_2 - \theta_3^2}{2\theta_3^2 y} - \frac{\theta_1 y}{2}$$

Black-Scholes-Merton:

$$F(x) = \frac{\log(x)}{\theta_2}$$

$$\mu_Y(y; \theta) = \frac{\theta_1}{\theta_2} - \frac{\theta_2}{2}$$

Capítulo 4

Ecuaciones Diferenciales Estocásticas y los Tipos de Interés

4.1. Modelo de Ait-Sahalia

En esta sección veremos una de las principales aplicaciones de las ecuaciones diferenciales estocásticas a la práctica financiera. Así, si consideramos, por ejemplo, un instrumento financiero como el bono cupón cero¹. Uno puede pensar que un bono cupón con vencimiento en el tiempo T , podría ser calculable mediante la siguiente esperanza bajo la medida neutral al riesgo:

$$B(0, T) = \mathbb{E}^{\mathbb{Q}}[\exp - \int_0^T r_u du] \quad 4.33 \quad (4.1)$$

donde el proceso del tipo corto de interés r_u podría ser cualquier proceso. En las sucesivas secciones, se presentarán algunos de los procesos estocásticos más difundidos, estudiados y ampliamente utilizados, como solución a la ecuación diferencial estocástica general siguiente

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \quad (4.2)$$

Numerosos procesos en tiempo continuo han sido propuestos para capturar el tipo de interés instantáneo o «spot». En esta sección, presentaremos el modelo de Ait-Sahalia² que satisface una EDE no-lineal del tipo:

$$dX_t = (\alpha_{-1}(X_t)^{-1} + \alpha_0 + \alpha_1 X_t + \alpha_2 (X_t)^2)dt + \beta_1 (X_t)^\rho dW_t. \quad (4.3)$$

Este modelo fue propuesto por Ait-Sahalia para modelar los tipos de interés y se trata de

¹«Fixed Income Mathematics: Analytical and Statistical Techniques», Fabozzi, F. J., 2006.

²«Testing Continuous-Time Models of the Spot Interest Rate», Ait-Sahalia, Yacine, (1996).

una generalización de los más famosos modelos paramétricos empleados en la literatura a la hora de capturar la dinámica de los tipos de interés «spot». Recordemos brevemente algunos de los conceptos y modelos que se han analizado en secciones precedentes.

En finanzas, un modelo en tiempo continuo usualmente se basa en uno o más procesos de difusión estacionarios $W_t, t \geq 0$, con dinámicas guiadas por ecuaciones diferenciales estocásticas de Itô:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (4.4)$$

donde $W_t, t \geq 0$ es un movimiento Browniano estándar. Las funciones $\mu(\Delta)$ y $\sigma^2(\Delta)$ son funciones de tendencia y difusión del proceso que, la mayoría de las ocasiones, están parametrizadas como:

$$\mu(x) = \mu(x, \delta) \quad \text{y} \quad \sigma^2(x) = \sigma^2(x, \delta)$$

donde $\delta \in \Theta \subset R^k$.

Supongamos que estamos ante una ecuación diferencial estocástica no-lineal del tipo

$$dX_t = (\alpha_{-1}(X_t)^{-1} + \alpha_0 + \alpha_1 X_t + \alpha_2 (X_t)^2)dt + \beta_1 (X_t)^\rho dW_t. \quad (4.5)$$

Que es satisfecha por un proceso en tiempo continuo como es el modelo propuesto por Aït-Sahalia. Este modelo paramétrico, ya visto en anteriores secciones, generaliza los más famosos modelos paramétricos usados en la literatura matemático-financiera a la hora de imitar la dinámica de los tipos de interés «spot».

Creo conveniente efectuar un breve recortorio, aún a riesgo de caer en la reiteración, sobre cómo se llega a los diferentes parámetros de tendencia y de difusión en los múltiple modelos de tasas spots a lo largo de la literatura. Así, aunque la mayoría de los modelos especifican un parámetro de tendencia lineal y de reversión a la media, $\mu(x, \delta) = \beta(\alpha - x)$, donde el tipo de interés spot se mueve en la dirección de la media α a una velocidad β ; sin embargo, a la hora de especificar el parámetro de difusión, la variación es muy destacable. Así:

- Si $\sigma^2(x, \delta) = \sigma^2$ estaríamos ante el modelo de Vasicek (1977).
- En cambio, si $\sigma^2(x, \delta) = \sigma^2 x$ ya el modelo especificado sería el de Cox, Ingersoll y Ross (1985), también denominado como modelo CIR.
- Aunque, si $\sigma^2(x, \delta) = \sigma^2 x^2$ entonces definiríamos el modelo de Courtdon (1982).

- En el modelo de Chan et al. (1992), el parámetro de difusión viene dado por $\sigma^2(x, \delta) = \sigma^2 x^\gamma$

- Si $\sigma^2(x, \delta) = \sqrt{\sigma_0 + \sigma_1 x}$ nos encontraríamos con el modelo de Duffie y Kan (1993).

En general, el modelo de Vasicek o de Ornstein-Uhlenbeck, el modelo CIR o modelo de la raíz cuadrada y otros como el de Merton, el de Black-Scholes-Merton, el de Brennan y Schwartz entre otros, que, aunque no destacan especialmente en la modelización de los tipos de interés ³, también se caracterizan por un parámetro lineal de tendencia, conforman familia de modelos ampliamente estudiada conocida como la familia Chan-Karolyi-Longstaff-Sanders (CKLS), también ya vista en secciones precedentes.

Los procesos CKLS son solución para la EDE parametrizada de la siguiente manera:

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 X_t^{\theta_4} dW_t. \quad (4.6)$$

donde $\theta_1 = \beta\alpha$ y $\theta_2 = -\beta X_t$, de tal forma que:

$$\mu(x) = \mu(x, \delta) = \beta(\alpha - x) = (\beta\alpha) - (\beta X_t) = (\theta_1 - \theta_2 X_t).$$

Bajo diferentes especificaciones paramétricas del modelo CLKS ⁴, obtendríamos algunos de los modelos mencionados con anterioridad, así:

- Si $\theta_4 = 0$, el proceso sería conducido por el modelo de Vasicek.
- Si $\theta_4 = 1/2$, el proceso sería guiado por el modelo CIR.
- Pero, si $\theta_4 = 1$, el proceso estaría especificado por el modelo de Brennan y Schwartz.
- Si $\theta_4 = 1$ y, además, $\theta_1 = 0$, el proceso estaría definido por el modelo de Black, Scholes y Merton.
- Y si $\theta_1 = 0$, el proceso sería llevado por el modelo de CEV.

Vemos que esta familia CKLS es una extensión de todos estos modelos y todos ellos con el parámetro de tendencia lineal. Está claro que el modelo CKLS es sólo un caso particular del modelo de Ait-Sahalia.

Frente a estos modelos, se han desarrollado otros que contemplan un parámetro de tendencia fuertemente no-lineal; así, surgieron:

³El modelo de Black-Scholes-Merton se usa ampliamente en modelar la dinámica de los precios de los derivados financieros, donde ha adquirido una notoria relevancia en el mundo financiero.

⁴En todos los casos, $\theta_3 > 0$.

• El modelo de Ahn-Gao que es una transformación del modelo CIR. El proceso es la solución a la EDE:

$$dX_t = X_t(\theta_2 - ((\theta_3)^3 - \theta_1\theta_2)X_t)dt + \theta_3(X_t)^{\frac{3}{2}}dW_t. \quad (4.7)$$

• El proceso Radial Ornstein-Uhlenbeck que es la solución a la EDE:

$$dX_t = (\theta(X_t)^{-1} - X_t)dt + dW_t \quad (4.8)$$

Como se observa en esta última ecuación, el parámetro de difusión no sólo es constante sino que es igual a la unidad, lo cual, aunque coyuntural o circunstancialmente pudiese arrojar unos buenos resultados, esta fuerte restricción lo podría descartar como modelo de uso habitual que capture la dinámica de los tipos de interés.

En las siguientes subsecciones estudiaremos el caso práctico de modelar una serie financiera de 2548 observaciones de tipos de interés «spot» y analizaremos las diferentes metodologías más ampliamente utilizadas y su problemática. Así, comenzaremos el estudio de la aplicabilidad de las ecuaciones diferenciales estocásticas el proceso de Ornstein-Uhlenbeck, para continuar seguidamente con el proceso CIR, ambos procesos son casos particulares del modelo de Ait-Sahalia. Finalmente, concluiremos con un modelo derivado del CIR, que denominaremos *CIR#*.

4.1.1. Modelo de Vasicek o de Ornstein-Uhlenbeck

Si $2\theta\beta > \sigma^2$, el proceso es estrictamente positivo, en otro caso, es no-negativo. Esto se traduciría en que podría alcanzar el estado 0; no obstante, esta posibilidad, en el caso de modelización de los tipos de interés, no fue admitida hasta el año 2009, cuando debido a la política monetaria del BCE que fijaba los tipos de interés próximos a cero, los tipos de interés alcanzaban valores negativos.

La siguiente parametrización es la más arraigada en el ámbito financiero

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t. \quad (4.9)$$

Particularmente, si pensamos en tipos de interés, un modelo de uso común es el modelo de Vasicek (o modelo de Ornstein-Uhlenbeck). A continuación, presentaremos una de las más

consideradas soluciones a la ecuación diferencial estocástica que sigue

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t. \quad (4.10)$$

se reduciría a

$$dX_t = (\alpha_0 + \alpha_1 X_t)dt + \beta_1 (X_t)^\rho dW_t.$$

la cual se podría reexpresar como:

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 X_t^{\theta_4} dW_t. \quad (4.11)$$

En primer lugar, se abordará el proceso de Ornstein-Uhlenbeck o proceso de Vasicek como solución única a la EDE siguiente

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 dW_t, \quad (4.12)$$

donde $\theta_1, \theta_2, \theta_3 \in \mathbb{R}_+$ y donde $X_0 = x_0 > 0$; que, como podemos observar, se trata de una particularidad de la ecuación 4.11, cuando $\theta_4 = 0$.

La ecuación, en el contexto financiero, se reexpresa como

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \quad (4.13)$$

donde $\sigma \in \mathbb{R}_+$ y $\mu, \theta \in \mathbb{R}$; con $X_0 = x_0 > 0$. A σ , se le el nombre de volatilidad, a μ , el valor de equilibrio del proceso en el largo plazo y a θ , la velocidad de reversión a la media, μ .

Así, podemos ver que, como aplicación del lema de Itô, la solución analítica para 4.12 y para 4.13, respectivamente, es

$$X_t = \frac{\theta_1}{\theta_2} + \left(x_0 - \frac{\theta_1}{\theta_2}\right)e^{-\theta_2 t} + \theta_3 \int_0^t e^{-\theta_2(t-u)} dW_u \quad (4.14)$$

si se refiere a 4.12, o si lo hace a 4.13 es

$$X_t = \mu + (x_0 - \mu)e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-u)} dW_u. \quad (4.15)$$

Resulta fácil ver que la relación entre una y otra ecuación es

$$\begin{aligned}\mu &= \frac{\theta_1}{\theta_2} \\ \theta &= \theta_2\end{aligned}$$

Si hacemos que $\theta_1 = 0$ ⁵, entonces el proceso es básicamente una exponencial negativa perturbada por la integral estocástica y quedaría como sigue

$$dX_t = -\theta_2 X_t + \theta_3 dW_t. \quad (4.16)$$

Cuya solución explícita sería

$$X_t = x_0 e^{-\theta_2 t} + \theta_3 \int_0^t e^{-\theta_2(t-u)} dW_u. \quad (4.17)$$

En un ámbito financiero, si únicamente pensamos en tipos de interés, la dinámica del modelo es como sigue

$$dr_t = k(\theta - r_t)dt + \beta dW_t^{\mathbb{Q}} \quad (4.18)$$

donde k , θ y β son constantes. El movimiento aleatorio es generado por el movimiento Browniano de medida Q , $W_t^{\mathbb{Q}}$. Una propiedad importante del proceso de Vasicek es que el tipo de interés revierte a la media, θ , y la tendencia de reversión está controlada por k . Asimismo, este proceso es un proceso de difusión y, por ende, markoviano, lo que conlleva a unas sencillas y atractivas fórmulas cerradas. En finanzas, la reversión a la media describe una serie temporal en la que «los rendimientos pueden ser muy inestables en el corto plazo pero muy estables en el largo plazo». Esto se podría expresar de otra manera diciendo que la desviación típica de los rendimientos anuales medios decrecen más rápido que la inversa del período de análisis. Esto tendría una serie de implicaciones, por ejemplo, que el proceso no es un paseo aleatorio, sino que períodos de rendimientos decrecientes son seguidos por retornos crecientes.

Finalmente, comentar que el valor futuro del tipo de interés está normalmente distribuido

⁵Originalmente propuesto por Ornstein y Uhlenbeck en su «On the Theory of Brownian Motion» (1930), *Phys. Rev.*, **36**, 823-841.

con la distribución

$$r_T \sim N\left(\theta(\rho_0 - \theta)e^{-kT}, \frac{\beta^2}{2k}(1 - e^{-2kT})\right).$$

El comportamiento de $r(t, T)$ puede ser cóncavo, convexo, etc. También podría ser negativa para alguno de los valores de los parámetros.

Para hacernos una idea de a qué se asemeja este proceso, seguidamente, se proyecta gráficamente

```
set.seed(1234567)
r0 <- 0.1013
theta <- 0.0023
k <- 0.2610
beta <- 0.4249
n <- 100 # Pruebas de simulación
T <- 1
m <- 252
dt <- T/m
r <- matrix(0,m+1,n)
r[1,] <- r0
t <- seq(0, T, dt)

for (j in 1:n){
  for(i in 2:(m+1)){
    dr <- k * (theta - r[i-1,j]) * dt + beta * sqrt(dt) * rnorm(1,0,1)
    r[i,j] <- r[i-1,j] + dr
  }
}

rT.expected <- theta + (r0 - theta) * exp(-k * t)
rT.stdev <- sqrt(beta^2/(2 * k) * (1 - exp(-2 * k * t)))
matplot(t, r[,1:10], type="l", lty=1, main= "Trayectorias Tasa de Interés", ylab="rt")
abline(h=theta, col="red", lty=2)
lines(t, rT.expected,lty=2)
lines(t, rT.expected + 2 * rT.stdev,lty=2)
```

```
lines(t, rT.expected - 2 * rT.stdev,lty=2)
points(0,r0)
```

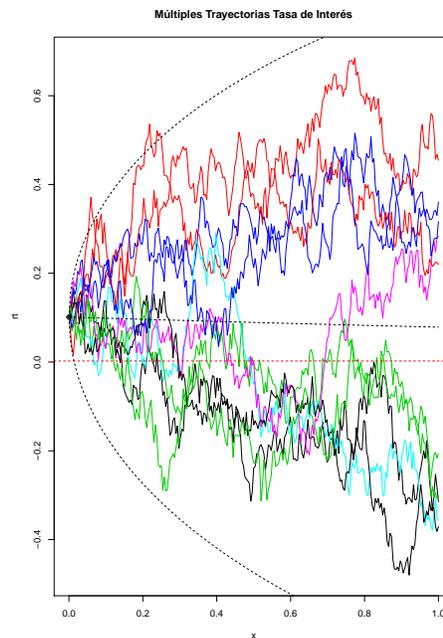


Figura 4.1: Las múltiples trayectorias simuladas de los tipos de interés por el método de aproximación de Euler.

En la gráfica 4.1, se puede observar que el proceso se inicia en 0.1013 que se corresponde con el valor inicial del tipo de interés, r_0 . También se puede percibir que, en el medio plazo, el proceso tiende a moverse hacia 0.023, esto es, θ , a una velocidad de 0.2610, es decir, k , la velocidad de reversión a la media.

Se han añadido unas líneas punteadas para destacar esto, así como una línea para el valor

esperado del proceso⁶ y unas bandas de confianza igual a $\mp 2\sqrt{v(t, x)}$, donde

$$v(t, x) = \text{Var}_\theta(X_t | X_0 = x_0) = \frac{\theta_3 \exp\{2(1 - \exp\{-\theta_2 t\})\}}{2\theta_2},$$

es decir, se ha considerado dos veces la desviación típica de la densidad condicional o de transición Gaussiana del proceso de Ornstein-Uhlenbeck.

No obstante, lo más destacable es la presencia de tipos de interés negativos en algunas de las trayectorias, donde el proceso cae por debajo del 0%. Hasta el año 2009, en un contexto financiero, una coyuntura de tipos de interés próximos a cero o incluso negativos era inadmisibles; no obstante, con la crisis de deuda soberana de Grecia y otros países europeos, las autoridades monetarias europeas decidieron bajar los tipos de interés a un nivel tal que la negatividad de los tipos de interés entró a formar parte de lo posible. Por lo que, a día de hoy, en la coyuntura actual de tipos de interés, el modelo de Ornstein-Uhlenbeck aparentemente conlleva una vigencia que nunca antes había disfrutado.

La muestra de datos que se usará en lo sucesivo es una serie financiera o temporal que denominaremos como EONIA. Los datos de EONIA (Euro Overnight Index Average) de nuestra serie se recogen desde el 23 de febrero de 2009 al 20 de febrero de 2019. EONIA contiene 2.548 valores de tipos de interés diarios⁷.

Con el objeto de obtener explícitamente la solución del tipo 4.14 a la ecuación diferencial estocástica 4.12, se necesita estimar el valor de los diferentes parámetros que la componen. Como ya se ha comentado en el capítulo 3, existen diferentes métodos de estimación de los parámetros teóricos, ya sean directa, como el método de inferencia por máxima verosimilitud, o indirectamente, como el de Euler, el de Elerian, etc. A continuación, se procederá a la aplicación de las diferentes metodologías más abundantemente empleadas en el mundo matemático-financiero.

Comenzaremos por el método de máxima verosimilitud. Sabemos que, bajo un conjunto

⁶Se correspondería a la media de la densidad de la distribución de X_t dado $X_0 = x_0$, que sería:

$$\mu(t, x) = \mathbb{E}_0(X_t | X_0 = x_0) = \frac{\theta_1}{\theta_2} + \left(x_0 - \frac{\theta_1}{\theta_2}\right) \exp\{-\theta_2 t\}$$

⁷Obtenida de BSmarkets.com. Informar que, en el mes de octubre del año 2019, esta serie será sustituida por **ESTER** (Euro Short-Term Rate); información extraída del diario **Expansión**, bajo el título «El BCE publicará en octubre de 2019 ESTER, la tasa que sustituirá al EONIA», 28, junio, 2018.

de asunciones: crecimiento lineal, global de Lipschitz, positividad del coeficiente de difusión, la existencia de los momentos hasta un orden determinado, etc., son condiciones suficientes para la consistencia y normalidad asintótica de los estimadores de máxima verosimilitud. Recordemos que si el parámetro θ_2 de la ecuación diferencial estocástica 4.12 es estrictamente positivo, el proceso de O-U es ergódico y posee una densidad estacionaria explícita con su media 2.4 y su varianza condicional 2.5.

Para la obtención de los estimadores de los parámetros por máxima verosimilitud, se implementará el siguiente código usando la muestra de datos EONIA

```
#####
# Con todos los datos de EONIA #
#####
require(stats4)
require(sde)
#--- Densidad condicional ---#
dcOU=function(x,t,x0,theta,log=FALSE){
  mc <- theta[1]/theta[2]+(x0-theta[1]/theta[2])*exp(-theta[2]*t)
  dc <- sqrt(theta[3]^2*(1-exp(-2*theta[2]*t))/(2*theta[2]))
  dnorm(x, mean = mc, sd = dc, log=log)
}
#--- Función de Verosimilitud ---#
OU.lik=function(theta1,theta2,theta3){
  n=length(X)
  dt=deltat(X)
  -sum(dcOU(x=X[2:n],t=dt,x0=X[1:(n-1)],theta=c(theta1,theta2,theta3),log=TRUE))
}
X <- Eonia[1:2548]
length(X)
which(is.na(X))
X <- na.omit(X)
which(is.na(X))
#--- Estimación de los Parámetros por Máxima Verosimilitud ---#
```

```

fit_tot_OU <- mle(OU.lik,start=list(theta1=0.01,theta2=0.01,theta3=0.01),
                 method="L-BFGS-B",lower=c(0.001,0.001,0.001),upper=c(1,1,1))

#--- Informe del Ajuste (Estimaciones y Errores Típicos) ---#
summary(fit_tot_OU)
# Coefficients:
#           Estimate Std. Error
# theta1 0.00100000 0.001567018
# theta2 0.01739983 0.003612560
# theta3 0.07626190 0.001082403

# -2 log L: -5861.512

```

Su ejecución nos ha arrojado unos estimadores de θ_1 , θ_2 y θ_3 que presentan unos errores estándares muy bajos, por debajo del 0.3%. Además, $\theta_3 > 0$ y $\theta_2 > 0$, lo cual el proceso es ergódico con densidad condicional con distribución Gaussiana.

Por tanto, la solución explícita única a la ecuación diferencial estocástica 4.12 es

$$X_t = 0.057472 + (x_0 - 0.057472)e^{-0.01739983t} + 0.07626190 \int_0^t e^{-0.01739983(t-u)} dW_u \quad (4.19)$$

La cual, al hablar de un proceso de tipos de interés, se podría reescribir como

$$r_t = 0.057472 + (r_0 - 0.057472)e^{-0.01739983t} + 0.07626190 \int_0^t e^{-0.01739983(t-u)} dW_u \quad (4.20)$$

Otros métodos de estimación paramétrica pueden ser aplicados para comprobar si los resultados que arrojan son similares o no a los obtenidos por máxima verosimilitud.

Otro método de obtención de los estimadores de los parámetros del proceso de Ornstein-Uhlenbeck consiste en el cálculo explícito de los mismos a través de unas ecuaciones formuladas por Bibby, Jacobsen & Sorensens (2005).

Para el cálculo exacto de \hat{k}_n , $\hat{\theta}_n$ y $\hat{\sigma}_n^2$, estimadores de k_n , θ_n y σ_n , parámetros de la ecuación

diferencial estocástica

$$dX_t = k(\theta - X_t)dt + \sigma dW_t, \quad (4.21)$$

donde $k, \theta \in \mathbb{R}$ y $\sigma \in \mathbb{R}_+$ es constante; recurriríamos a las fórmulas explícitas siguientes

$$\hat{k}_n = -\ln \left(\frac{\frac{1}{(n-1)} (\sum_{i=2}^n r_{i-1}) (\sum_{i=2}^n r_i) - \sum_{i=2}^n r_{i-1} r_i}{\frac{1}{(n-1)} (\sum_{i=2}^n r_{i-1})^2 - (\sum_{i=2}^n (r_{i-1})^2)} \right), \quad (4.22)$$

$$\hat{\theta}_n = \frac{1}{1 - e^{-\hat{k}_n}} \left(\frac{1}{(n-1)} \sum_{i=2}^n r_i - \frac{e^{-\hat{k}_n}}{(n-1)} \sum_{i=2}^n r_{i-1} \right), \quad (4.23)$$

$$\hat{\sigma}_n^2 = \frac{2\hat{k}_n}{1 - e^{-2\hat{k}_n}} \cdot \frac{1}{(n-1)} \sum_{i=2}^n (r_i - r_{i-1} e^{-\hat{\theta}_n} - \hat{\theta}_n (1 - e^{-\hat{k}_n}))^2. \quad (4.24)$$

Con $e^{-\hat{k}_n} > 0$, condición necesaria para la existencia de las ecuaciones 4.22, 4.23 y 4.24.

Así, al implementar el código que, a continuación, se expone se obtendrían los valores de los estimadores de los parámetros del proceso de O-U solución a la ecuación diferencial estocástica 4.21 que mejor ajusta la muestra EONIA

```
##### ----- Estimadores Parám. Ornstein-Uhlenbeck por Bibby et al ----- #####
### --- Estimadores  $\hat{\theta}_1$ ,  $\hat{\theta}_2$  y  $\hat{\theta}_3$  --- ###
### ---  $\hat{k}$  ---
Eo <- precios[1:2547]
n <- length(Eo)
s1 = 0
for (i in 2:n){
  s1 = s1 + Eo[i]/Eo[i-1]
}
A1 <- (n-1)*s1
s2 = 0
s3 = 0
for(i in 2:n){
  s2 = s2 + Eo[i]
```

```
}
for(i in 2:n){
  s3 = s3 + 1/Eo[i-1]
}
A2 <- s2*s3
s4 = 0
s5 = 0
for(i in 2:n){
  s4 = s4 + Eo[i-1]
}
for(i in 2:n){
  s5 = s5 + 1/Eo[i-1]
}
A3 <- s4*s5
khat <- -log((A1 - A2)/((n-1)^2 - A3))

###---  $\hat{\theta}$ 
Eo <- precios[1:2547]
n = length(Eo)
s1 = 0
for (i in 2:n){
  s1 = s1 + Eo[i]
}
A1 <- 1/(n-1)*s1
A2 <- exp(-khat)
A3 <- (1 - A2)*(n-1)
thetahat <- A1 + (Eo[2547]-Eo[1])*A2/A3

### ---  $\hat{\sigma}$ 
Eo <- precios[1:2547]
n = length(Eo)
s1 = 0
```

```

for (i in 2:n){
  s1 = s1 + (1/Eo[i-1])*(Eo[i] - Eo[i-1]*exp(-khat) - thetahat*(1-exp(-khat)))^2
}
A1 <- s1
s2 = 0
for (i in 2:n){
  s2 = s2 + (1/Eo[i-1])*(((thetahat/2 - Eo[i-1])*exp(-2*khat) -
(thetahat-Eo[i-1])*exp(-khat) + thetahat/2)/khat)
}
A2 <- s2
sigmahat <- sqrt(A1/A2)
estBibby <- c(khat,thetahat,sigmahat)
estBibby
# 0.007719985 0.186241878 0.092754586

```

No obstante, para que las comparaciones sean homogéneas, necesitamos reparametrizar los resultados obtenidos en términos de la EDE 4.12, con lo que los estimadores son $\hat{\theta}_1 = 0.001437784$, $\hat{\theta}_2 = 0.007719985$ y $\hat{\theta}_3 = 0.092754586$.

Así, con estos estimadores, la solución explícita de O-U a la ecuación 4.12 es

$$r_t = 0.186242 + (r_0 - 0.186242)e^{-0.007719985t} + 0.092754586 \int_0^t e^{-0.007719985(t-u)} dW_u \quad (4.25)$$

Finalmente, los parámetros también podrían ser estimados mediante un enfoque diferente e interesante para el caso de un proceso de difusión con volatilidad instantánea constante, como es el caso de un proceso de Ornstein-Uhlenbeck o de Vasicek puro. No obstante, no se debería restringir su aplicabilidad a este supuesto, pues, si a cualquier proceso estocástico se le aplica la transformación de Lamperti, ya estudiada en una sección anterior, el catálogo de aplicabilidad de la metodología que se va a presentar se amplía sustancialmente.

El modelo de Ornstein-Uhlenbeck se podría reexpresar en términos discretos como una EDE de la manera que se plasma a continuación. Sea $\{W_{t \geq 0}\}$ un movimiento Browniano definido sobre un espacio fijo de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y sea $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ su filtración aumentada, que representa la información del mercado hasta el tiempo t .

El tipo de interés «spot» r_t es conducido por la ecuación diferencial estocástica

$$dr_t = \theta(\mu - r_t)dt + \sigma dW_t. \quad (4.26)$$

Donde k es la velocidad de reversión a la media, θ es la media a la que revierte el proceso y σ es la volatilidad instantánea que es constante.

Si se expresa $\beta_0 = k\theta$ y $\beta_1 = 1 - k$, entonces la ecuación diferencial estocástica 4.26 se reexpresaría como un proceso autorregresivo de orden uno con tendencia de la forma

$$r_t = \beta_0 + \beta_1 r_{t-1} + \varepsilon_t \quad (4.27)$$

Donde $\{\varepsilon_t\}$ son variables aleatorias independientes que siguen una distribución normal $\mathcal{N}(0, \sigma^2)$ ⁸

Por ende, mediante un modelo *ARIMA*(1, 1, 1) se estiman los parámetros del coeficiente drift del proceso de Ornstein-Uhlenbeck.

Los resultados de la estimación de los parámetros β_0 y β_1 del modelo empleado 4.27 son como sigue

$$r_t = 0.0006 + 0.7376r_{t-1}.$$

Por tanto, podríamos calcular los parámetros que conforman el coeficiente drift de la EDE

$$dr_t = \theta(\mu - r_t)dt + \sigma dW_t,$$

donde $\theta = 1 - \beta_1$ y $\mu = \frac{\beta_0}{\theta}$.

Así, $\theta = 0.2610$ y $\mu = 0.0023$ que, recordemos, son la velocidad de reversión a la media y

⁸Donde su media incondicional es

$$\mathbb{E}[r_t] = \frac{\beta_0}{(1 - \beta_1)}$$

y su varianza incondicional es

$$Var[r_t] = \frac{\sigma^2}{(1 - \beta_1^2)} = \frac{\sigma^2}{(1 - (1 - \theta^2))}.$$

Su varianza condicionada a r_{t-1} es

$$Var[r_t | r_{t-1}] = \sigma^2$$

la media a la que revierte el proceso de difusión, respectivamente.

Que, homogeneizando la parametrización, los estimadores de los parámetros θ_1 y θ_2 componentes del coeficiente drift son $\hat{\theta}_1 = 0.0006$ y $\hat{\theta}_2 = 0.0023$.

Y para la obtención del estimador del parámetro difusión θ_3 , calculamos la varianza condicional, σ^2 y, a través de la ecuación $\sqrt{Var[r_t]} = \sqrt{\frac{\sigma^2}{(1-\beta_1^2)}}$, ya se tendría que $\hat{\theta}_3 = 0.62926$.

Así, la solución Ornstein-Uhlenbeck a la ecuación diferencial estocástica 4.12 es

$$r_t = 0.00261 + (r_0 - 0.00261)e^{-0.0023t} + 0.62926 \int_0^t e^{-0.0023(t-u)} dW_u \quad (4.28)$$

Por tanto, a través de los diferentes de estimación paramétrica, se han conseguido las soluciones explícitas siguientes:

Por el método de aproximación por máxima verosimilitud

$$r_t = 0.057472 + (r_0 - 0.057472)e^{-0.01739983t} + 0.07626190 \int_0^t e^{-0.01739983(t-u)} dW_u \quad (4.29)$$

Por el método de Bibby et al

$$r_t = 0.186242 + (r_0 - 0.186242)e^{-0.007719985t} + 0.092754586 \int_0^t e^{-0.007719985(t-u)} dW_u \quad (4.30)$$

Por el método autorregresivo de orden uno

$$r_t = 0.00261 + (r_0 - 0.00261)e^{-0.0023t} + 0.62926 \int_0^t e^{-0.0023(t-u)} dW_u \quad (4.31)$$

con $r_0 = 0.1013$.

Se puede observar que, al contrastar los resultados de las diversas metodologías empleadas, los resultados más cercanos se logran para los estimadores de los parámetros θ_2 y θ_3 bajo el método de máxima verosimilitud y de Bibby et al.

Debo añadir que al emplear toda la muestra para la estimación de los parámetros teóricos del proceso de O-U, claramente estamos incurriendo en una suposición «a priori» falsa en lo que respecta a la constancia del parámetro difusión, θ_3 . Se debería proceder a efectuar un análisis para comprobar si éste se mantiene constante o no. En el supuesto de variación en la

volatilidad, se debería detectar los puntos donde suceden estos cambios y proceder a estimar los parámetros en cada una de las submuestras resultantes inmediatamente antes del punto de cambio. Esta metodología será analizada con detalle en secciones posteriores.

Supongamos que se desconociese o ignorase la solución explícita del proceso de Vasicek a la EDE, y se procediese a la simulación de la trayectoria del proceso solución a la EDE,

$$dr_t = \theta(\mu - r_t)dt + \sigma dW_t,$$

emplearíamos una serie de métodos de aproximación o de discretización de la solución continua a la EDE. En el presente estudio, usaremos, específicamente para el proceso de O-U, los métodos de Euler y método de Ozaki.

En primer lugar, supongamos que usemos la EDE con la siguiente parametrización, aunque se podría usar otra parametrización cualquiera:

$$dr_t = (\theta_1 - \theta_2 r_t)dt + \theta_3 dW_t,$$

con un valor inicial $R_{t_0} = R < -0$ y con $0 = t_0 < t_1 < \dots < t_N = T$, donde N es el número de intervalos en el que se divide t . Consideremos el proceso de Ornstein-Uhlenbeck solución de esa ecuación diferencial estocástica. Entonces, la aproximación de Euler de R es un proceso estocástico continuo:

$$R_{*_{i+1}} = R_{*_{i}} + b(t_i, R_{*_{i}})(t_{i+1} - t_i) + \sigma(t_i, R_{*_{i}})(W_{i+1} - W_i),$$

con $t \in [t_i, t_{i+1})$.

Supongamos que los valores de los parámetros inicialmente sean

$$r_0 = 0.1013 \quad \theta = 0.0023 \quad k = 0.2610 \quad \beta = 0.4249$$

Mediante el esquema de discretización de Euler, se arroja el siguiente resultado

```
require(sde)
set.seed(1234)
drift <- expression(0.0023*(0.2610-x))
difussion <- expression(0.4249)
X1 <- sde.sim(X0=0.1013,drift=drift,sigma=difussion)
plot(X1, main="Aproximación Euler de Proceso O-U")
```



Figura 4.2: Trayectoria simulada del proceso de Ornstein-Uhlenbeck mediante el esquema de discretización de Euler cuyo funcionamiento es bueno siempre que Δ sea pequeño; con un Δ grande, presenta dificultades, llegando incluso a explotar.

como se puede deducir al ver el proceso estocástico R^* , la simulación de éste no es más que efectuar una simple simulación del incremento del proceso de Wiener,

$$(W_{i+1} - W_i).$$

Así, para el caso que nos ocupa, $b(t, r) = (\theta_1 - \theta_2 r)$ y $\sigma(t, r) = \theta_3$, se fija un $R_0 = r$ y se toma como $\vec{\theta} = (\theta_1, \theta_2, \theta_3)$ y $dW_t = \sqrt{\Delta t}Z$, donde $Z = rnorm(0, 1)$. El siguiente código con el subsiguiente gráfico donde se muestra la trayectoria simulada del proceso de O-U mediante el método de Euler por una simple simulación del incremento del movimiento Browniano

```
set.seed(1234)
N <- 252
T <- 1
```

```
x <- 0.1013
theta <- c(0.0023, 0.2610, 0.4249)
dt <- 1/N
Y <- numeric(N+1)
Y[1] <- x
Z <- rnorm(N)
for (i in 1:N)
  Y[i+1] <- Y[i] + theta[1]*(theta[2]-Y[i])*dt + theta[3]*sqrt(dt)*Z[i]
Y <- ts(Y, start=0, deltat=dt)
plot(Y, main="set.seed(1234)")
N <- 252
T <- 1
x <- 0.1013
theta <- c(0.0023, 0.2610, 0.4249)
dt <- 1/N
Y <- numeric(N+1)
Y[1] <- x
Z <- rnorm(N)
for (i in 1:N)
  Y[i+1] <- Y[i] + theta[1]*(theta[2]-Y[i])*dt + theta[3]*sqrt(dt)*Z[i]
Y <- ts(Y, start=0, deltat=dt)
plot(Y, main="Aproximación Euler de Proceso 0-U")
```

Los críticos del método de Euler esgrimen que únicamente genera soluciones a ecuaciones diferenciales determinísticas, esto es, es una aproximación de una simple constante, con lo que su abanico de aplicabilidad se reduce bastante.

Sin embargo, existe otros esquemas de aproximación cuya idea principal descansa en el principio que una aproximación lineal es bastante mejor que una aproximación determinística. Así pues, se podría emplear un método que aproxime localmente el drift con una función lineal, el método Ozaki que funciona bien bajo la suposición de una difusión constante, es decir, funciona bien con EDEs homogéneas u homocedásticas. Lo cual es supuestamente un método adecuado para aplicar en el proceso de Ornstein-Uhlenbeck. Así, mediante el método

Ozaki se obtendría la siguiente ecuación diferencial

$$r_{t+\Delta t} = r_t + \frac{b(r_t)}{b_r(r_t)} \left(\exp\{b_r(r_t)\Delta t - 1\} \right).^9 \quad (4.32)$$

Por el uso de un esquema de Ozaki, un método de linealización local para una EDE homogénea como la de nuestro ejemplo, se obtiene

```
set.seed(12345)
drift <- expression(0.0023*(0.2610-x))
difussion <- expression(0.4249)
sde.sim.ozaki <- function(X0=0.1013, t0, dt, N=100, d1=drift, d1.x, s1=difussion){
  X <- numeric(N+1)
  B <- function(x) d1(1, x)
  Bx <- function(x) d1.x(1, x)
  X[1] -> X0
  vd <- all.vars(drift)
  vs <- all.vars(sigma)
  if(length(vd)!=1 | length(vs)>0)
    stop("drift must depend on 'x' and volatility must be constant")
  if((length(vd) == 1) && (vd != "x"))
    stop("drift must depend on 'x'")
  for(i in 2:(N+1)){
    x <- X[i-1]
    Kx <- log(1+B(x)*(exp(B(x)*dt)-1)/(x*Bx(x)))/dt
    Ex <- x + B(x)/Bx(x)*(exp(Bx(x)*dt)-1)
    Vx <- S^2 * (exp(2*Kx*dt)-1)/(2*Kx)
    X[i] <- rnorm(1, mean=Ex, sd=sqrt(Vx))
  }
}
```

⁹Donde

$$b(r_t) = \frac{dr_t}{dt}$$

y

$$b_r(r_t) \frac{dr_t}{dt} = \frac{d^2 r_t}{dt^2},$$

son las derivadas de primer y segundo orden respecto de t .

```
X
}
plot(X, main="Aproximación OZAKI PROCESO O-U")
```

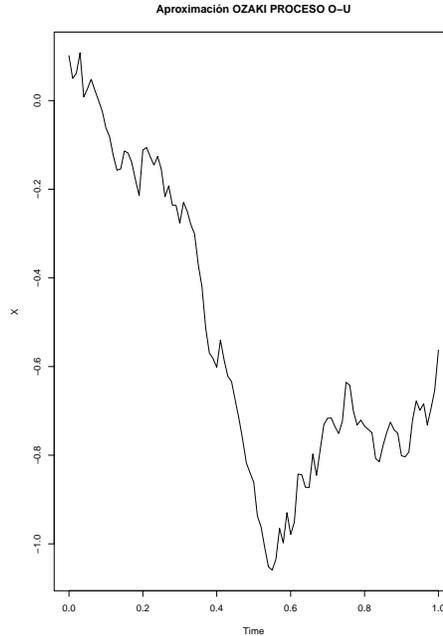


Figura 4.3: Trayectoria del proceso de Ornstein-Uhlenbeck cuando se aproxima mediante el método de linealización de Ozaki. Éste representa el método idóneo para la simulación de soluciones a ecuaciones diferenciales estocásticas homogéneas. Si la EDE posee un coeficiente drift lineal y una volatilidad instantánea constante, los métodos de Euler, Ozaki y Shoji-Ozaki arrojan los mismos resultados.

Mediante la implementación de un sencillo código, obtendríamos la simulación de los incrementos del proceso, donde $R[i + 1] \sim N(E_r, V_r)$ y $R[i] = r$ ¹⁰ y

$$V_r = \sigma^2 \frac{\exp\{2K_r \Delta t\} - 1}{2K_r}.$$

¹⁰Donde:

$$E_r = r + \frac{b(r)}{b_r(r)} \left(\exp\{b_r(r) \Delta t\} - 1 \right).$$

La constante K_t se calcula así:

$$K_t = \frac{1}{\Delta t} \log \left(1 + \frac{b(R_t)}{R_t b_r(R_t)} (\exp(b_r(R_t) \Delta t) - 1) \right).$$

Y, finalmente, si se aplica el método de linealización local de Shoji-Ozaki, su trayectoria sería

```

set.seed(123456)
sde.sim.shoji <- function(X0=0.1013, t0, dt, N=100, d1, d1.x, d1.xx, d1.t, s1){
  X <- numeric(N+1)
  S <- s1(1, 1)
  X[1] -> X0
  for(i in 2:(N+1)){
    x <- X[i-1]
    Lx <- d1.x(dt, x)
    Mx <- S^2 * d1.xx(dt, x)/2 + d1.t(dt, x)
    Ex <- (x + d1(dt, x)*(exp(Lx*dt)-1)/Lx + Mx*(exp(Lx*dt)-1-Lx*dt)/Lx^2)
    Vx <- S^2 * (exp(2*Lx*dt)-1)/(2*Lx)
    X[i] <- rnorm(1, mean=Ex, sd=sqrt(Vx))
  }
  X
}
plot(X, main="Aproximación SHOJI-OZAKI PROCESO O-U")

```

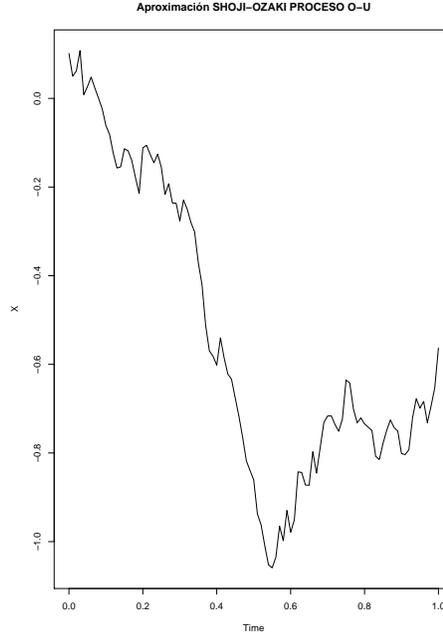


Figura 4.4: Trayectoria simulada del proceso de Ornstein-Uhlenbeck cuando se aplica el método de linealización de Shoji-Ozaki.

Como se muestra, la trayectoria resultante es idéntica, pues, en este caso concreto, la equivalencia entre la aproximación de Euler con Ozaki y Shoji-Ozaki es máxima; así, un esquema de Euler sobre una ecuación diferencial estocástica con la volatilidad constante arroja el mismo resultado que un esquema de Ozaki o un método de Shoji-Ozaki sobre la misma EDE.

Por tanto, disponemos de una solución exacta y de una solución simulada a través de diferentes métodos de aproximación; aunque insistimos que para la discretización se podría haber utilizado cualquier otro método que encajase en la naturaleza del proceso de Ornstein-Uhlenbeck solución a la EDE sea cual sea la parametrización de la misma.

A la altura del análisis en la que nos encontramos, debemos recordar cómo se podría calcular el precio de un bono cupón cero; así, a través de la expresión

$$B(0, T) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_0^T r_u du} \right], \quad (4.33)$$

se obtendría el valor de un bono cupón cero $B(0, T)$ con vencimiento en el tiempo, T . En la ecuación 4.33 anteriormente reflejada, interviene muy activamente un proceso de tipo de interés, r_u , que podría estar conducido por una EDE 4.2 cuya solución podría estar determinada por otro proceso estocástico, como por ejemplo, el proceso de Ornstein-Uhlenbeck o proceso de Vasicek.¹¹

Para calcular el precio del bono cupón cero, se podría estimar esa esperanza usando una simulación de Monte Carlo, pero el modelo de Vasicek nos permite computar el valor del bono cupón cero al hacer uso de la propiedad de Markov y las técnicas de ecuaciones diferenciales parciales. Usando este método, se puede preciar el bono mediante el uso de las siguientes ecuaciones:

$$B(0, T) = e^{-a(T)-b(T)r_0} \quad (4.34)$$

donde

$$b(\tau) = \frac{1 - e^{-\tau\kappa}}{\kappa} \quad (4.35)$$

y

$$a(\tau) = \left(\theta - \frac{\beta^2}{2\kappa}\right) (\tau - b(\tau)) \frac{\beta^2}{4\kappa} b(\tau)^2 \quad (4.36)$$

Esta solución cerrada para un bono cupón cero hace que nuestra vida sea más fácil puesto que no se necesita calcular la esperanza para obtener el precio de un bono.

Así, si pensamos en el bono cupón cero como $B(0, T) = e^{-y_T T}$ donde y_T es el rendimiento en tiempo continuo para un bono cupón cero con vencimiento T , y, mediante una serie de arreglos, llegaríamos a una expresión de la forma $y_T = \frac{-\ln B(0, T)}{T}$.

Asimismo, mediante esta última forma de expresar el rendimiento de un bono cupón cero en tiempo continuo, se podría llegar a la obtención de la curva de rendimientos, representación ésta altamente apreciada en el mundo financiero.

Así, tomando en consideración estas expresiones, vamos a desarrollar el código R que

¹¹La literatura financiero-matemática habitualmente denomina como proceso de Vasicek al proceso de Ornstein-Uhlenbeck con el parámetro θ_1 igual a cero; aunque, en nuestro caso, procurando rendir homenaje indistintamente a los tres reputados personajes, nos referiremos al proceso por cualquiera de las dos denominaciones indiferenciadamente.

implementa la función que devuelve el precio exacto bajo el modelo de Vasicek.

```
r0 <- 0.1013
k <- 0.2610
theta <- 0.0023
beta <- 0.4249
T <- 1

precioVasicekBCC <- function(r0, k, theta, beta, T){
  b.vas <- (1/k)*(1-exp(-T*k))
  a.vas <- (theta-beta^2/(2*k^2))*(T-b.vas)+(beta^2)/(4*k)*b.vas^2
  return(exp(-a.vas-b.vas*r0))
}
prVasicek_BC0 <- exp(-a.vas-b.vas*r0)
prVasicek_BC0
# 0.9374615
```

Así, el precio **Vasicek Bono Cupón Cero** calculado, $PrExactoVas = 0.9374615$.

Como ya se había anticipado en esta sección más arriba, la curva de rendimiento puede ser creciente o decreciente con el vencimiento, T ; así, usaremos la función de precios anterior para generar formas varias de la mencionada curva (véase gráfica 4.5).

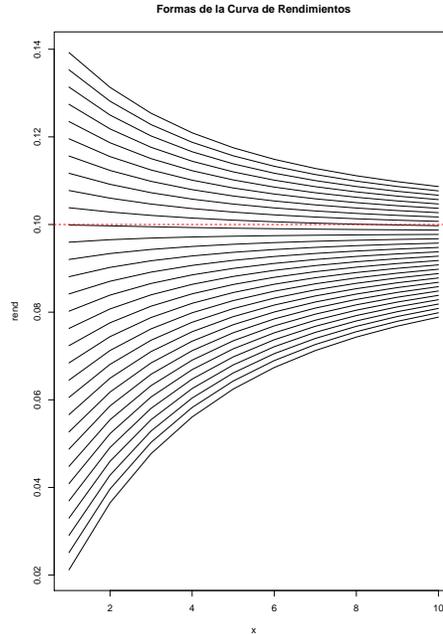


Figura 4.5: Diferentes formas de la curva de rendimientos: creciente o decreciente con el vencimiento.

Finalmente, para cerciorarnos de que la expresión del precio del bono cupón cero se ha implementado con corrección, se procede a comparar el precio usando esta fórmula frente al precio usando una simulación de Monte Carlo con el objeto de estimar la esperanza $\mathbb{E}^{\mathbb{Q}}$, bajo la medida \mathbb{Q} .

```

r0 <- 0.1013
k <- 0.2610
theta <- 0.0023
beta <- 0.4249
n <- 1000 # Número de pruebas MC
T <- 1
m <- 252
dt <- T/m

```

```

r <- matrix(0, m+1, n)
r[1,] <- r0
for(j in 1:n){
  for(i in 2:(m+1)){
    dr <- k*(theta-r[i-1,j])*dt + beta*sqrt(dt)*rnorm(1, 0, 1) # Método Euler
    r[i, j] <- r[i-1, j] + dr
  }
}

ss <- colSums(r[2:(m+1),]*dt) # Integral
c <- exp(-ss)
estimador <- mean(c)
stdErr <- sd(c)/sqrt(n)
Vasicekprecioexacto <- preciovasicekBCC(r0, k, theta, beta, T)

cat('Precio Exacto con Vasicek:',round(Vasicekprecioexacto,4),'n')
cat('Precio Monte Carlo:',round(estimador,4),'n')
cat('Error Estándar Monte Carlo:',round(stdErr,4),'n')

```

Se presentan, a continuación, los resultados del código anterior

PrecioExactoconVasicek : 0.9375

PrecioMonteCarlo : 0.9373

ErrorEstándarMonteCarlo : 0.0068

Los precios son muy similares, existe una diferencia relativa porcentual del 0.02134%, levemente más alta que la obtenida por la discretización de Euler cuyo un error es del 0.0068%.

4.1.2. Modelo de Cox-Ingersoll-Ross (CIR)

Como ya hemos mencionado en la sección anterior, los tipos de interés son fundamentales para los mercados financieros, el consumo privado, las ganancias empresariales o corporativas, los precios de los activos, la inflación y la economía. Adicionalmente, hemos visto gráficamente

que, a medida que los tipos de interés alcanzan valores altos, también se aprecia un incremento en su volatilidad¹². La dinámica de los tipos de interés es muy diferente a la de otras variables, como los rendimientos de las acciones, por ejemplo. Hasta finales de la primera década del presente siglo, hablar de tipos de interés negativos era impensable. Desde aquellos años a esta parte, las políticas monetarias expansivas implementadas por las autoridades monetarias europeas rompieron el techo de lo imperante en la teoría económico-financiera, de hecho los tipos de interés habían alcanzado valores negativos. Adicionalmente, la heterocedasticidad y la regresión a la media, son propiedades de los tipos de interés; a medida que los tipos de interés suben por encima del nivel de la media, hay un drift que «tira» de los tipos hacia abajo, y, viceversa, cuando los tipos de interés caen por debajo de la media, el drift tira de los tipos hacia arriba. Para capturar estas dos características especialmente relevantes de los tipos de interés, Cox, Ingersoll y Ross construyeron un modelo.

El modelo que se va a analizar en esta sección nace como una supuesta mejora del proceso de Vasicek al restringir que los tipos de interés sólo puedan tomar valores positivos o, al menos, no-negativos, dependiendo del cumplimiento o no de ciertas condiciones paramétricas que mencionaremos más adelante. El modelo de Cox-Ingersoll-Ross, desde ahora será referido como CIR, es una particularización de la familia de modelos de Chan-Karolyi-Longstaff-Sanders que es una clase de ecuaciones diferenciales estocásticas que son solución a la EDE:

$$dX_t = (\theta_1 + \theta_2 X_t)dt + \theta_3 X_t^{\theta_4} dW_t. \quad (4.37)$$

Que si $\theta_4 = 1/2$, se expresaría como:

$$dX_t = (\theta_1 + \theta_2 X_t)dt + \theta_3 X_t^{1/2} dW_t. \quad (4.38)$$

Que es el modelo CIR, el cual, si $2\theta_1 > \theta_3^2$ y $\theta_2 < 0$, toma valores en $(0, +\infty)$, si no, en $[0, +\infty)$. La ecuación anterior se puede reparametrizar como sigue:

$$dr_t = k(\alpha - r_t)dt + \sigma r_t^{1/2} dW_t. \quad (4.39)$$

Que, como se puede observar, el drift es lineal e igual al de modelo de Vasicek, que, recordemos,

¹²Ver gráfica 4.6.

era:

$$dr_t = k(\alpha - r_t)dt + \sigma dW_t, \quad (4.40)$$

que ignora la heterocedasticidad de la volatilidad.

El parámetro drift es lineal y exactamente igual al de la EDE del proceso de Ornstein-Uhlenbeck. Si el parámetro θ_4 de la ecuación 4.37 es cero estaríamos ante el proceso de O-U. Los parámetros θ_1 , θ_2 y θ_3 de la ecuación 4.37, son constantes.

$W_{t \geq 0}$ es un movimiento Browniano definido sobre $(\Omega, \mathcal{F}, \mathbb{P})$ que se trata de un espacio fijo de probabilidad con $\mathbb{F} = \mathcal{F}_{t \geq 0}$, su filtración aumentada.

En la ecuación 4.37, figura el parámetro drift que ya hemos dicho que es lineal, y el parámetro difusión, aleatorio o estocástico con una varianza $\theta_3^2 X_t^{2\theta_4}$, lo cual está en sintonía con lo mencionado pocas líneas arriba, esto es, a medida que se incrementan los tipos de interés, la varianza o la volatilidad asimismo sube.

Para ver la aplicabilidad del proceso CIR en la modelización del tipo de interés y, por tanto, su mayor o menor utilidad en un contexto económico-financiero, se va a probar ante una base de datos, que hemos denominado serie financiera EONIA, que se inicia en febrero del 2009 y finaliza en febrero del 2019; contiene un total de 2548 observaciones diarias de los tipos de interés «spot» conocidos en el mercado de tipos como 'overnight spot rates'. La elección de la serie temporal no resulta de una mera decisión azarosa o arbitraria, sino que obedece a un criterio muy pensado, pues el período que abarca la serie representa un giro inesperado en el mercado de los tipos de interés y, por ende, la sorprendente aparición de valores negativos en los tipos de interés spot.

Así, EONIA representada gráficamente sería como sigue

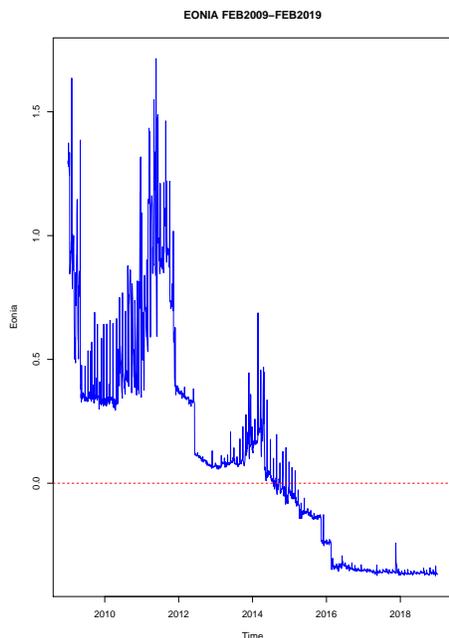


Figura 4.6: Serie Eonia de tipos de interés spot que contiene 2548 observaciones desde el 23 de febrero de 2009 hasta el 20 de febrero de 2019.

Como se puede observar en el gráfico 4.6, los tipos de interés alcanzan valores negativos desde la primera mitad del año 2014, hecho éste sin precedentes hasta la fecha en un contexto económico-financiero.

Por ello, analizar si el proceso CIR se ajusta bien o no a los datos de nuestra serie, se podría anticipar exitosamente con carácter inmediato, dada la presencia de valores negativos en la serie EONIA. No se debe olvidar, que el proceso CIR surge como mejora al proceso de Vasicek, debido a que el modelo de Ornstein-Uhlenbeck admitía valores negativos o cero, y esto, contrastaba con la teoría económico-financiera vigente en esas fechas.

Aún así, recordemos que el proceso CIR es solución a una ecuación diferencial estocástica del tipo

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 \sqrt{X_t} dW_t, \quad (4.41)$$

con θ_1, θ_2 y $\theta_3 \in \mathbb{R}_+$. Asimismo, sabemos que su solución explícita es 4.14.

Entonces, si aplicamos un método de estimación de los parámetros por verosimilitud exacta, se necesitaría implementar el código siguiente (uso de la versión reescalada exponencialmente de la función modificada de Bessel de primera clase de orden q , $I_q(\cdot)$)

```
#--- Función de Bessel Reescalada Exponencialmente ---#
expBes <- function(x,nu){
  mu <- 4*nu^2
  A1 <- 1
  A2 <- A1*(mu- 1)/(1*(8*x))
  A3 <- A2*(mu- 9)/(2*(8*x))
  A4 <- A3*(mu-25)/(3*(8*x))
  A5 <- A4*(mu-49)/(4*(8*x))
  A6 <- A5*(mu-81)/(5*(8*x))
  A7 <- A6*(mu-121)/(6*(8*x))
  (1/sqrt(2*pi*x))*(A1-A2+A3-A4+A5-A6+A7) }

#--- Densidad condicional ---#
dcCIR <- function(x,t,x0,theta,log=FALSE){
  c <- 2*theta[2]/((1-exp(-theta[2]*t))*theta[3]^2)
  ncp <- 2*c*x0*exp(-theta[2]*t)
  df <- 4*theta[1]/theta[3]^2
  u <- c*x0*exp(-theta[2]*t)
  v <- c*x
  q <- 2*theta[1]/theta[3]^2-1
  lik <- log(c)-(u+v)+(q/2)*log(v/u)+log(expBes(2*sqrt(u*v),q))+2*sqrt(u*v)
  if (!log){lik=exp(lik)}
  lik
}

#--- Función de Verosimilitud ---#
CIR.lik <- function(theta1,theta2,theta3){
  n <- length(X)
```

```

dt <- deltat(X)
-sum(dcCIR(x=X[2:n],t=dt,x0=X[1:(n-1)]),theta=c(theta1,theta2,theta3),log=TRUE))
}

```

```

library(sde)
X <- Eonia[1:2548]
X <- na.omit(X)
which(is.na(X))
#--- Estimación de los Parámetros por Máxima Verosimilitud ---#
fit_tot <- mle(CIR.lik,start=list(theta1=0.1,theta2=0.1,theta3=0.1),
              method="L-BFGS-B",lower=c(0.001,0.001,0.001),upper=c(1,1,1))

# Error in optim(start, f, method = method, hessian = TRUE, ...) :
# L-BFGS-B needs finite values of 'fn'

```

Vemos que no funciona porque la aproximación tiende a explotar. Sabemos que la densidad transitoria del proceso CIR, si existe, sigue una distribución χ^2 no-centrada. Por lo cual, podemos hacer uso de la vía χ^2 de aproximación; así, el código anterior sufriría unas leves modificaciones y quedaría como sigue

```

#--- Densidad condicional ---#
dcCIR2 <- function(x,t,x0,theta,log=FALSE){
  c <- 2*theta[2]/((1-exp(-theta[2]*t))*theta[3]^2)
  ncp <- 2*c*x0*exp(-theta[2]*t)
  df <- 4*theta[1]/theta[3]^2
  lik <- (dchisq(2*x*c, df=df, ncp=ncp, log=TRUE))
  if (!log){lik=exp(lik)}
  lik
}

#--- Función de Verosimilitud ---#

```

```

CIR.lik2 <- function(theta1,theta2,theta3){
  n <- length(X)
  dt <- deltat(X)
  -sum(dcCIR2(x=X[2:n],t=dt,x0=X[1:(n-1)],theta=c(theta1,theta2,theta3),log=TRUE))
}

library(sde)
X <- Eonia[1:2548]
X <- na.omit(X)
which(is.na(X))

#--- Estimación de los Parámetros por Máxima Verosimilitud ---#
fit_tot_chi <- mle(CIR.lik2,start=list(theta1=0.1,theta2=0.1,theta3=0.1),
  method="L-BFGS-B",lower=c(0.001,0.001,0.001),upper=c(1,1,1))

# Error in optim(start, f, method = method, hessian = TRUE, ...) :
# L-BFGS-B needs finite values of 'fn'

```

El código tampoco funciona, también explota. Estos hechos entrarían dentro de lo posible pues, tal y como hemos estudiado en el capítulo 2, la densidad transitoria de un proceso CIR sigue una distribución χ^2 que, aunque sea una distribución de probabilidad muy conocida, no deja de ser computacionalmente difícil.

Recordemos que la densidad condicional resulta ser

$$p_{\theta}(t, y|x_0) = ce^{-(u+v)} \left(\frac{u}{v}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}),$$

donde $I_q(\cdot)$ es una función Bessel de primera clase de orden q , y donde, $u = \frac{2\theta_2}{(\theta_3^2(1-e^{-\theta_2 t}))} x_0 e^{-\theta_2 t}$ y $v = \frac{2\theta_2}{(\theta_3^2(1-e^{-\theta_2 t}))} y$.

También somos conocedores que los grados de libertad de la distribución ν dependen de los parámetros θ_1 , velocidad de reversión a la media, y θ_3 , la volatilidad instantánea; por ende, cuando $t \rightarrow 0$, entonces u y v explotan. Y esto podría suceder cuando la volatilidad es muy baja y la velocidad de reversión es alta, i.e. en periodos de tipos de interés bajos, los grados de libertad se incrementan, puesto que $\nu = \frac{4\theta_1}{\theta_3^2}$ y esto conllevaría evaluar la χ^2 en la cola de

la derecha.

El problema es muy parecido con la función Bessel reescalada exponencialmente; la χ^2 está basada en la función Bessel, $I_q(\cdot)$, y cuando ésta tiene que ser evaluada en la coyuntura anterior también explota.

Una vez explicada la posible causa del no-funcionamiento de los anteriores códigos, una de las preguntas que se nos viene a la mente sería ¿es el método de inferencia de los estimadores por máxima verosimilitud no-válido? La respuesta es no.

Si en lugar de inferir usando toda la muestra EONIA, únicamente utilizamos un fragmento de la misma, los resultados que se obtienen son como siguen

```
### --- Para la primera submuestra, Eonia[1:85] --- ###
### Función Bessel Reescalada Exponencialmente
expBes=function(x,nu){
  mu=4*nu^2
  A1=1
  A2=A1*(mu- 1)/(1*(8*x))
  A3=A2*(mu- 9)/(2*(8*x))
  A4=A3*(mu-25)/(3*(8*x))
  A5=A4*(mu-49)/(4*(8*x))
  A6=A5*(mu-81)/(5*(8*x))
  A7=A6*(mu-121)/(6*(8*x))
  (1/sqrt(2*pi*x))*(A1-A2+A3-A4+A5-A6+A7) }

#--- Densidad Condicional ---#
dcCIR=function(x,t,x0,theta,log=FALSE){
  c=2*theta[2]/((1-exp(-theta[2]*t))*theta[3]^2)
  ncp=2*c*x0*exp(-theta[2]*t)
  df=4*theta[1]/theta[3]^2
  u=c*x0*exp(-theta[2]*t)
  v=c*x
  q=2*theta[1]/theta[3]^2-1
  lik=log(c)-(u+v)+(q/2)*log(v/u)+log(expBes(2*sqrt(u*v),q))+2*sqrt(u*v)
```

```
if (!log){lik=exp(lik)}
lik }

#--- Función de Verosimilitud ---#
CIR.lik=function(theta1,theta2,theta3){
  n=length(X)
  dt=deltat(X)
  -sum(dcCIR(x=X[2:n],t=dt,x0=X[1:(n-1)]),theta=c(theta1,theta2,theta3),log=TRUE)) }

X <- Eonia[1:85] # Un fragmento de la muestra EONIA
X <- na.omit(X)
plot(X)
which(X<0)
is.na(X)
#--- Estima los Parámetros por Máxima Verosimilitud ---#
fit_1=mle(CIR.lik,start=list(theta1=0.15,theta2=0.15,theta3=0.15),
          method="L-BFGS-B",lower=c(0.001,0.001,0.001),upper=c(1,1,1) )

#--- Informe del Ajuste (Estimaciones y Errores Típicos) ---#
summary(fit_1)
# Coefficients:
#           Estimate Std. Error
# theta1 0.1733261 0.06801607
# theta2 0.1934661 0.07861057
# theta3 0.1618493 0.01379620

# -2 log L: -96.41743
```

Vemos que se pueden obtener los estimadores por máxima verosimilitud de los parámetros teóricos de un proceso CIR solución continua a una ecuación diferencial estocástica. Otros fragmentos de la muestra original arrojan estimaciones de los parámetros del modelo CIR.

Por otro lado, si observamos con atención la representación gráfica de Eonia, se percibe que la serie parece que «salta», posiblemente como consecuencia de las intervenciones de las autoridades monetarias europeas. Parece que la volatilidad de los tipos, θ_3 , varía con intensidad diferente si se consideran los años 2016, 2017 y 2018, que durante el período que va del 2010 al 2012 (fenómeno del clustering de volatilidad: en períodos de volatilidad alta, son seguidos por volatilidad alta, y viceversa, a t de baja volatilidad, les siguen períodos de baja volatilidad).

Esto me lleva inicialmente a concluir que, aunque sea posible obtener una estimación de los verdaderos parámetros del modelo CIR, su efectividad, a la hora de valorar un activo financiero, un derivado o cualquier otro instrumento financiero, estará seriamente afectada.

A pesar de las dificultades anteriores, las estimaciones de los parámetros θ_1 , θ_2 y θ_3 , se pueden obtener explícitamente. Así, según Bibby et al, se podrían calcular los estimadores de los parámetros teóricos como sigue

```
### --- Estimadores  $\hat{\theta}_1$ ,  $\hat{\theta}_2$  y  $\hat{\theta}_3$  --- ###
### ---  $\hat{k}$ 
Eo <- precios[1:2547]
s1 = 0
for (i in 2:n){
  s1 = s1 + Eo[i]/Eo[i-1]
}
A1 <- (n-1)*s1

s2 = 0
s3 = 0
for(i in 2:n){
  s2 = s2 + Eo[i]
}

for(i in 2:n){
  s3 = s3 + 1/Eo[i-1]
}
```

```
A2 <- s2*s3

s4 = 0
s5 = 0
for(i in 2:n){
  s4 = s4 + Eo[i-1]
}
for(i in 2:n){
  s5 = s5 + 1/Eo[i-1]
}
A3 <- s4*s5

khat <- -log((A1 - A2)/((n-1)^2 - A3))

###---  $\hat{\theta}$ 
Eo <- precios[1:2547]
n = length(Eo)
s1 = 0
for (i in 2:n){
  s1 = s1 + Eo[i]
}
A1 <- 1/(n-1)*s1
A2 <- exp(-khat)
A3 <- (1 - A2)*(n-1)

thetahat <- A1 + (Eo[2547]-Eo[1])*A2/A3

### ---  $\hat{\sigma}$ 
Eo <- precios[1:2547]
n = length(Eo)
s1 = 0
for (i in 2:n){
```

	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$
Eonia	-0.002234969	-0.015599200	0.100029997

Tabla 4.1: Estimadores de los parámetros teóricos del proceso de Cox-Ingersoll-Ross para la muestra Eonia (23-02-2009/20-02-2019).

```

s1 = s1 + (1/Eo[i-1])*(Eo[i] - Eo[i-1]*exp(-khat) - thetahat*(1-exp(-khat)))^2
}
A1 <- s1
s2 = 0
for (i in 2:n){
  s2 = s2 + (1/Eo[i-1])*(((thetahat/2 - Eo[i-1])*exp(-2*khat) -
(thetahat-Eo[i-1])*exp(-khat) + thetahat/2)/khat)
}
A2 <- s2

sigmahat <- sqrt(A1/A2)

### --- Reparametrización
theta_1hat <- khat*thetahat
theta_2hat <- khat
theta_3hat <- sigmahat
theta_hat <- c(theta_1hat, theta_2hat, theta_3hat)
theta_hat

```

Ya se han estimado los parámetros y se estaría en disposición de calcular explícitamente la solución a la ecuación diferencial estocástica 2.8, que como ya sabemos sería

$$X_t = \left(x_0 - \frac{\theta_1}{\theta_2}\right)e^{-\theta_2 t} + \theta_3 e^{-\theta_2 t} \int_0^t e^{\theta_2 u} \sqrt{X_u} dW_u. \quad (4.42)$$

Es decir, tenemos una EDE de la forma 2.8, que traducido a tipos de interés y con los parámetros anteriormente estimados, quedaría como sigue

$$dr_t = (-0.002234969 + 0.015599200r_t)dt + 0.100029997\sqrt{r_t}dW_t.$$

Y su solución, si $r_0 = 0.1013$, sería

$$r_t = -0.041975e^{0.015599200t} + 0.100029997 \int_0^t e^{0.015599200t} \sqrt{r_u} dW_u,$$

que, si deseamos predecir el valor estimado de los tipos de interés en $t = 2549$, entonces se debería resolver la ecuación anterior.

Antes de continuar, debo llamar la atención del lector sobre el siguiente hecho; como se observa θ_1 y θ_2 son negativos, lo cual no cumple con las restricciones paramétricas ya comentadas en su apartado. Adicionalmente, $2\theta_1 < \theta_3^2$, lo cual deriva en que el proceso sea no-negativo. Por último, simplemente comentar que si $e^{-\hat{k}} = 0.9977675$, estrictamente positivo, se cumple la condición de existencia de los estimadores de Bibby et al.

La cuestión que se nos plantea es: ¿es el proceso CIR inhábil para modelar la muestra de datos EONIA? La respuesta es no, y nos apoyaremos en el desarrollo de la siguiente sección para dar sustento a esta conclusión.

4.1.3. Modelo CIR#

Como ya se ha visto en la sección precedente, el proceso de difusión CIR parece inadecuado para capturar la estructura temporal de los tipos de interés pues, aparentemente, no modela la actual coyuntura económico-financiera de tipos de interés negativos. Asimismo, el término de difusión en la dinámica de los tipos, cuando éstos son bajos, tiende a cero. Adicionalmente, también se ha expuesto que tanto la media «a largo» como la volatilidad no cambian con el tiempo.

Por ello, se ha sugerido un nuevo modelo, el modelo *CIR#*, que pretende captar la estructura temporal de los tipos, preservando la estructura de volatilidad del mercado. Debido al destacado impacto que supuso la irrupción del modelo CIR como una importante mejora sobre el modelo de Ornstein-Uhlenbeck, estimo conveniente, dada su inextricable relación con el modelo *CIR#*, recordar las principales características del CIR.

Se trata de una generalización del modelo de O-U al caso de una volatilidad no constante y asume que la evolución del subyacente, los tipos de interés «a corto», es conducido por un

proceso de difusión. Es un proceso Markoviano, continuo, y solución única a la EDE¹³:

$$dr_t = [k(\theta - r_t) - \lambda(t, r_t)]dt + \sigma\sqrt{r_t}dW_t, \quad \text{con } r_0 > 0, \quad (4.43)$$

donde $W_{t \geq 0}$, un movimiento Browniano estándar.

La EDE 4.46 es un modelo homogéneo en el tiempo de un factor porque los parámetros k , θ y σ son independientes del tiempo, y la dinámica de los tipos está únicamente guiado por λr_t , el precio del riesgo del mercado $\lambda(t, r_t)$, con λ , invariable. Las trayectorias del proceso CIR nunca alcanzan valores negativos y su comportamiento depende de la relación de los tres parámetros, k , θ y σ , donde k, θ y $\sigma \geq 0$.

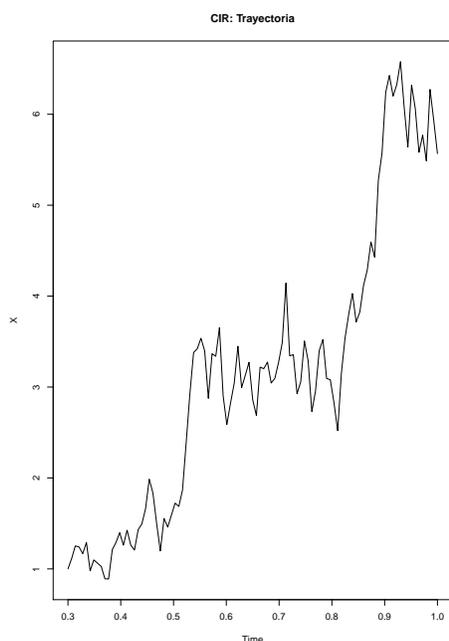


Figura 4.7: El gráfico muestra una trayectoria simple simulada de un proceso de difusión de Cox-Ingersoll-Ross que, como se puede observar, nunca toma valores negativos.

También cabe recordar que el proceso alcanza valores estrictamente positivos siempre que se cumpla la condición $2k\theta \geq \sigma^2$. Por tanto, los tipos negativos quedarían excluidos.

¹³Otra forma alternativa de expresar el proceso de Vasicek. En la literatura especializada, las EDEs se expresan y reexpresan de variadas formas dependiendo fundamentalmente el objeto de estudio.

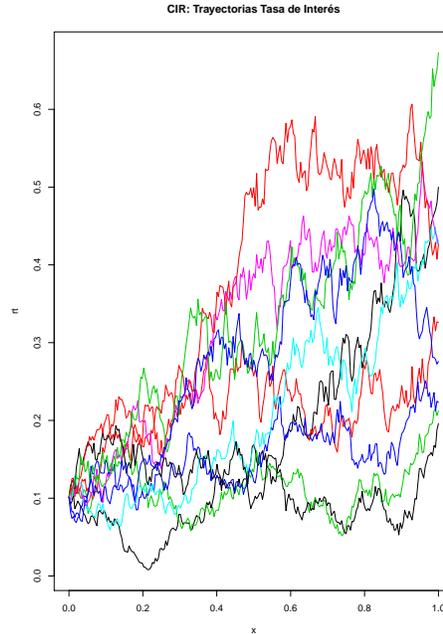


Figura 4.8: Si la condición $2k\theta \geq \sigma^2$, ó $2\theta_1 \geq \theta_3^2$, bajo otra parametrización, es necesaria para que el Proceso CIR sea estrictamente positivo; así, se puede percibir en el gráfico cómo todas las trayectorias son positivas.

Sin embargo, si se incumple la anterior condición, lo que equivale a decir que $2k\theta < \sigma^2$, los tipos de interés también podrían alcanzar el valor cero, como se puede ver en el gráfico [4.9](#).

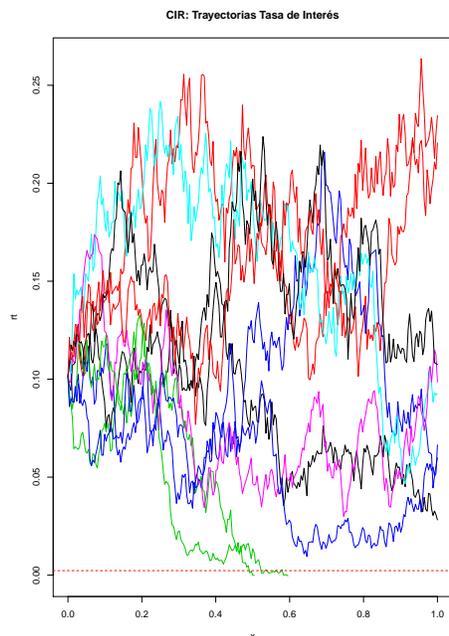


Figura 4.9: Si la condición $2k\theta \geq \sigma^2$, ó $2\theta_1 \geq \theta_3^2$, bajo otra parametrización, es necesaria para que el proceso CIR sea estrictamente positivo; en el caso de incumplimiento, el proceso es no-negativo lo que implica que puede alcanzar el estado 0; así, se puede ver claramente en el gráfico cómo dos trayectorias simuladas son absorbidas por el cero.

El término de difusión, $k(\theta - r_t)$ tiende a cero cuando r_t es bajo, lo cual circunstancialmente puede ser contradictorio con los datos de mercado.

El modelo CIR asume la no existencia de «saltos». Asunción poco realista, por otra parte, ya que los tipos de interés probablemente son una de las variables más intervenidas por las autoridades monetarias correspondientes. Así, esta consideración se olvida de las políticas monetarias y fiscales de los diferentes bancos centrales, políticas claramente intervencionistas que distorsionan artificialmente los mercados financieros con estos «shocks» o «innovaciones».

El modelo CIR, asimismo, asume que la volatilidad instantánea, σ , es constante.

De ahí, la necesidad de la presencia de un modelo que ajuste los tipos de interés, incluido los negativos, que tome en consideración esas intervenciones de las autoridades monetarias reflejadas en forma de saltos; un modelo que incluya una volatilidad condicional de los cambios

en los tipos de interés como no constante. De esta manera, nace el modelo $CIR\#^{14}$, con el objeto de preservar la estructura del modelo original CIR que describe las dinámicas de los tipos de interés observados en los mercados financieros, proporcionando una nueva metodología que intenta ofrecer respuestas a las diversas asunciones anteriormente destacadas.

Para ello, sería prioritario detectar estos «saltos» o cambios bruscos estadísticamente significativos de los tipos de interés de nuestra muestra. Una forma de ejecutarlo sería particionando la muestra en submuestras, no necesariamente del mismo tamaño, que logren capturar estos cambios significativos. Para llevar a cabo la partición de nuestra muestra, se podría efectuar arbitrariamente con un tamaño prefijado de submuestra, por ejemplo, $length = 200$, obteniendo así 12 submuestras de este tamaño y una más corta, de tamaño 148. No obstante, existen alternativos métodos de enfoque o aproximación, menos arbitrarios, que detectan los múltiples cambios en la distribución de probabilidad de un proceso estocástico o una serie financiera (serie de tiempo). En el caso que nos ocupa, se ha optado por una metodología paramétrica, otra no paramétrica y una tercera, en forma de algoritmo propuesta por Lavielle (2005).¹⁵

En cuanto a los dos primeros enfoques, se va a hacer uso de la función $cpoint()$ de R que básicamente identifica el instante, o bien, en tiempo discreto, $k0$, o bien, en tiempo continuo, $tau0$, en el que ocurre un cambio en el parámetro θ del proceso de difusión solución a la EDE

$$dr(t) = b(r(t))dt + \sigma(r(t))dW_t, \quad (4.44)$$

donde $b(\cdot)$ y $\sigma(\cdot)$ son funciones conocidas y $\theta \in \Theta \subset \mathbb{R}$ es el parámetro de estudio.

Así, dada una muestra de observaciones en el intervalo $[0, T = n\delta_n]$, se busca indentificar si ha habido un cambio significativo en el valor del parámetro, y cuándo esta modificación ha ocurrido; asimismo, estima consistentemente el valor del parámetro antes y después del punto de cambio. Stefano Maria Iacus, autor de la función anteriormente referida, por simplicidad, asume como $k0$ el instante en tiempo discreto en el que el cambio acaece, donde $k0 \in 1, \dots, n$. El método de estimación empleado es el de Mínimos Cuadrados. Para la obtención de los estimadores por Mínimos Cuadrados, aplica el método de discretización o aproximación de

¹⁴Versión del modelo CIR creada por Giuseppe Orlando, Rosa Maria Mininni y Michele Bufalo, de junio de 2018.

¹⁵Véase **Using Penalized Constrasts for the Change-Point Problem**, *Signal Processing*85(8), de M. Lavielle, 2005.

Euler. Huelga comentar que todas las hipótesis de este método de discretización son asumidas. En nuestro caso particular, se cumplen escrupulosamente, pues estamos en el proceso de difusión CIR que posee un parámetro drift lineal, *conditio sine qua non* de la aproximación de Euler.

El estimador del punto de cambio, \hat{k}_0 se obtiene como:

$$\hat{k}_0 = \min_k \left(\min_{\theta_1, \theta_2} \left\{ \sum_{i=1}^k (Z_i^2 - \theta_i^2)^2 + \sum_{i=k+1}^n (Z_i^2 - \theta_i^2)^2 \right\} \right),$$

donde $k \in \{2, \dots, n-1\}$

O bien,

$$\hat{k}_0 = \max_k |D_k|,$$

donde,

$$D_k = \frac{k}{n} - \frac{S_k}{S_n},$$

donde,

$$S_n = \sum_{i=1}^n Z_i^2, \quad S_k = \sum_{i=1}^k Z_i^2,$$

donde $k \in \{2, \dots, n-1\}$

A modo de representación implementada, asumiendo que tenemos un modelo CIR, como sigue

$$dr(t) = (-0.002234969 + 0.015599200r(t))dt + 0.100029997\sqrt{r(t)}dW_t$$

A continuación, se muestra la trayectoria única de dos pasos r_1 y r_2 de un proceso CIR, el primero, en $[0, \tau_0]$ y el segundo, en $[\tau_0, 1]$ y el gráfico del estadístico $|D_k|$:

```
### Código Trayectoria única de unión r1 y r2 MODELO CIR y GRÁFICO de estadístico |D_{k}|
## {Trayectoria Única Proceso CIR: obtenida de r1 y r2}
## {Gráfico del Estadístico \abs{D_{k}}}
```

```
tau0 <- 0.695
k0 <- ceiling(1000*tau0)
```

```
set.seed(12345)
r1 <- sde.sim(X0 = 1, N = 2*k0, t0 = 0, T = tau0, model = "CIR", theta = c(6, 2, 1))
r2 <- sde.sim(X0 = r1[2*k0 + 1], N = 2*(1000-k0), t0 = tau0, T = 1, model = "CIR",
theta = c(6, 2, 1))
Y <- ts(c(r1, r2[-1]), start = 0, deltat = deltat(r1))
X <- window(Y, deltat = 0.01)
DELTA <- deltat(X)
n <- length(X)

mu <- function(x) 6-2*x # Parámetro Drift
sigma <- function(x) sqrt(X) # Parámetro Difusión
ResEst <- (diff(X)-mu(X[1:n-1])*DELTA)/sqrt(DELTA)*sigma(X[1:n-1])
tau <- seq(0, 1, length = length(ResEst))
k <- ceiling(n*tau)
Sn <- cumsum(ResEst^2)
Sk <- sum(ResEst^2)
Dk <- abs(k/n - Sn/Sk)

k0 <- which(Dk == max(Dk))
tau[k0]

par(mar = c(3, 3, 1, 1))
par(mfrow = c(2, 1))

plot(X, main="Trayectoria Única Proceso CIR: obtenida de r1 y r2")
abline(v = tau0, col = "red", lty = 3)

plot(tau, Dk, type = "l", main="Estadístico |D_k|")
abline(v = tau[k0], col = "blue")
abline(v = tau0, col = "red", lty = 3)
```

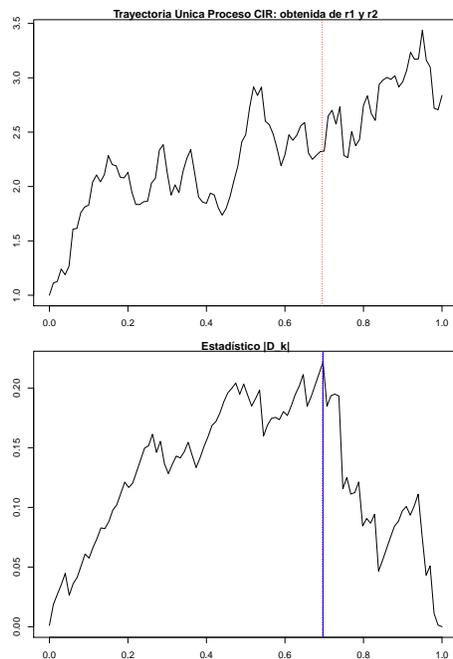


Figura 4.10: En el gráfico superior, se muestra cómo varía la volatilidad o parámetro θ_3 ó σ en el tiempo $\tau_0 = 0.695$ en el proceso simulado CIR. En el gráfico inferior, se representa la forma del estadístico $|D_k|$ y su máximo en el tiempo $\tau_0 = 0.695$. El punto de cambio, \hat{k}_0 , es la solución a $\hat{k}_0 = \operatorname{argmax}|D_k|$.

Seguidamente, de modo gráfico, se indican los diversos puntos de cambio en la serie financiera de EONIA, donde realmente existe una variación en el parámetro θ . Así:

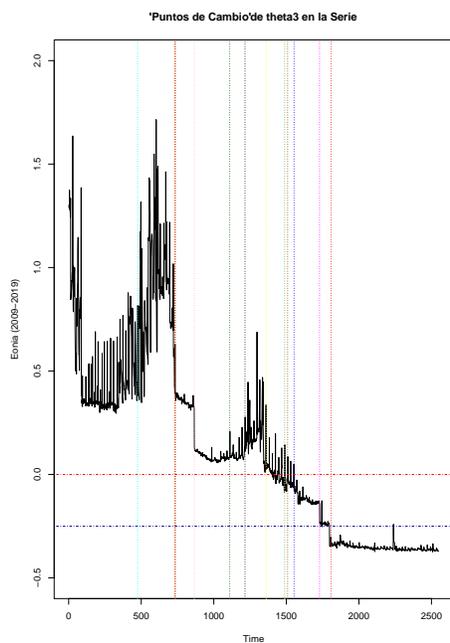


Figura 4.11: La gráfica representa los puntos de cambio del parámetro θ_3 , lo que sirve de inestimable ayuda para implementar el enfoque CIR# o cualquier otro.

Tras la identificación de estos saltos, saltos que posiblemente estén provocados por la implementación de cambios en las políticas monetarias, fiscales, etc., de las autoridades del BCE o de cualquier país miembro de la UE, o de cualquier otro banco central relevante como la Federal Reserve Bank, por ejemplo.

Debemos tener en cuenta el tamaño de la submuestra (cuanto más pequeño el tamaño, más fino el análisis) y la habilidad para capturar cualesquiera saltos (cuanto más grande el tamaño, mejor en términos de significancia estadística). En el caso de la partición de la muestra de EONIA, según los resultados arrojados por la función *cpoint()*, y a la vista de su representación gráfica, invita a afirmar que la división en grupos o submuestras parece muy razonable.

A continuación, se muestran en forma tabular (Tabla 4.2) los resultados obtenidos por ambas metodologías.

Así, si la partición de la muestra EONIA, se efectúa no paramétricamente, se obtienen los

	Eo1	Eo2	Eo3	Eo4	Eo5	Eo6	Eo7
<i>cpoint</i> ()	1 - 85	86 - 343	344 - 486	487 - 726	727 - 864	865 - 937	938 - 1214
<i>lavielle</i> ()	1 - 85	86 - 408	409 - 727	728 - 865	866 - 950	951 - 1214	1215 - 1296

	Eo8	Eo9	Eo10	Eo11	Eo12
<i>cpoint</i> ()	1215 - 1296	1297 - 1551	1552 - 1728	1729 - 2548	
<i>lavielle</i> ()	1297 - 1339	1340 - 1403	1404 - 1551	1552 - 1728	1729 - 2548

Tabla 4.2: Partición de la muestra EONIA en subconjuntos usando dos metodologías diferentes: una primera, *cpoint*(), no-paramétrica, y otra, *lavielle*(), basada en el algoritmo de Lavielle (2005).

siguientes 12 grupos 1–85, 86–343, 344–486, 487–726, 727–864, 865–937, 938–1214, 1215–1296, 1297–1551, 1552–1728, 1729–2548. Y si nos inclinamos por el algoritmo propuesto por Lavielle (2005), la muestra de EONIA se subdividiría en los 13 conjuntos que siguen: 1–85, 86–408, 409–727, 728–865, 866–950, 951–1214, 1215–1296, 1297–1339, 1340–1403, 1404–1551, 1552–1728, 1729–2548.

A continuación, se procedería a la estimación de los parámetros de los grupos de la muestra original particionada¹⁶.

Puesto que el modelo CIR no ajusta tipos de interés negativos e, igualmente, tampoco ajusta volatilidades normales o/y altas, cuando el valor del tipo de interés es bajo, los tipos de interés deberían ser trasladados de los valores cero o/y negativos. Para ello se puede llevar a cabo una transformación del tipo:

$$r_{trasl}(t) = r_{eonias}(t) + \alpha, \quad t \in [0, T],$$

donde $\alpha > 0$ es un término constante.

Mediante esta transformación se respeta la naturaleza estocástica de la dinámica del proceso, i.e. $dr_{trasl}(t) = dr_{eonias}(t)$. La elección del parámetro α se podría realizar arbitrariamente, no obstante, en nuestro análisis, seguiremos los consejos de G. Orlando, R. M. Mininni y M.

¹⁶Dado que el grupo 1297–1339, solamente tendría 43 valores, consideramos necesario recombinarlo con el siguiente conjunto 1340–1403, que contendría 64 observaciones, resultando así el nuevo grupo 1297–1403, con 107 datos.

	Eo1	Eo2	Eo3	Eo4	Eo5	Eo6
M. Armónica	0.8293438	0.3612351	0.7253829	0.3490567	0.1004379	0.07769209

	Eo7	Eo8	Eo9	Eo10	Eo11
M. Armónica	0.1630326	0.03753564	-0.09124331	-0.1132361	-0.3382219

Tabla 4.3: Media armónica de los diversos grupos procedentes de la partición de la muestra original EONIA.

Bufalo¹⁷ cuya opinión va dirigida a obtener el percentil 99 de la distribución condicional del proceso usando la muestra completa y, a pesar de ello, los valores negativos todavía persisten, optan por usar un α igual al percentil 1 de la distribución condicional del proceso CIR de los tipos de interés. Por ende, en este caso, la traslación sería:

$$r_{trasl}(t) = r_{eonia}(t) - \alpha, t \in [0, T].$$

Previamente, antes de hacer uso de la traslación, se procede al cálculo de la media armónica de cada uno de los grupos particionados de la muestra EONIA y se verifica si es más pequeña que un valor arbitrario, e. g. 0.01. En caso que esta condición se cumpla en al menos uno de los grupos, entonces se aplica la traslación a la muestra completa de EONIA.

Así, tras el cálculo de la media armónica de los correspondientes grupos, los conjuntos de datos denotados como $Eo9$, $Eo10$ y $Eo11$, arrojan unos valores de -0.09124331 , -0.1132361 y -0.3382219 , respectivamente, obviamente inferiores a 0.01; por ello, se procederá a transformar la muestra original, EONIA.

Para poder llevar a cabo la transformación arriba reseñada, se antoja necesario el cálculo de α . El valor que se obtiene es de -0.366 , correspondiente al percentil 1, como ya se ha comentado anteriormente. Así, la serie trasladada sería:

$$r_{trasl}(t) = r_{eonia}(t) + 0.366, \text{ con } t \in [0, T].$$

Tras la traslación de la serie original, nos adentramos en la fase crucial de nuestro estudio,

¹⁷ «On the Calibration of Short-Term Interest Rates through a CIR model», pág. 11 (12 de Junio de 2018).

esto es, el cálculo de los parámetros de la serie financiera eventualmente transformada $(+\alpha)$.

En este punto, estimamos conveniente recordar la EDE objeto de análisis:

$$dr_t = \tilde{k}(\tilde{\theta} - r_t) \cdot dt + \sigma\sqrt{r_t}dW_t, \quad \text{con } r_0 > 0, \quad (4.45)$$

donde $W_{t \geq 0}$, un movimiento Browniano estándar.

Que como podemos observar es una variación en la parametrización de la EDE:

$$dr_t = [k(\theta - r_t) - \lambda(t, r_t)]dt + \sigma\sqrt{r_t}dW_t, \quad \text{con } r_0 > 0, \quad (4.46)$$

donde $W_{t \geq 0}$, es un movimiento Browniano estándar.

Ambas ecuaciones diferenciales son exactamente iguales, simplemente estableciendo que:

$$k = \tilde{k} + \lambda > 0, \quad \theta = \frac{\tilde{k} \cdot \tilde{\theta}}{k} > 0.$$

Sabemos de la existencia de diferentes métodos de aproximación de los parámetros de los diversos procesos de difusión que son soluciones continuas a las ecuaciones diferenciales estocásticas. Entre el variado elenco, nos hemos decantado por el uso del método de aproximación de los estimadores por máxima verosimilitud.

Cabe recordar que, con el objeto de estimar por máxima verosimilitud, se aplica a cada una de las submuestras de la original EONIA el código siguiente de la función Bessel reescalada exponencialmente de la densidad transitoria

```
### --- Submuestra Eo1 --- ###
# --- Función Bessel Reescalada Exponencialmente --- #
expBes=function(x,nu){
  mu=4*nu^2
  A1=1
  A2=A1*(mu- 1)/(1*(8*x))
  A3=A2*(mu- 9)/(2*(8*x))
  A4=A3*(mu-25)/(3*(8*x))
  A5=A4*(mu-49)/(4*(8*x))
  A6=A5*(mu-81)/(5*(8*x))
  A7=A6*(mu-121)/(6*(8*x))
```

```

    (1/sqrt(2*pi*x))*(A1-A2+A3-A4+A5-A6+A7) }
#--- Densidad Condicional ---#
dcCIR=function(x,t,x0,theta,log=FALSE){
  c=2*theta[2]/((1-exp(-theta[2]*t))*theta[3]^2)
  ncp=2*c*x0*exp(-theta[2]*t)
  df=4*theta[1]/theta[3]^2
  u=c*x0*exp(-theta[2]*t)
  v=c*x
  q=2*theta[1]/theta[3]^2-1
  lik=log(c)-(u+v)+(q/2)*log(v/u)+log(expBes(2*sqrt(u*v),q))+2*sqrt(u*v)
  if (!log){lik=exp(lik)}
  lik }
#--- Función de Verosimilitud ---#
CIR.lik=function(theta1,theta2,theta3){
  n=length(X)
  dt=deltat(X)
  -sum(dcCIR(x=X[2:n],t=dt,x0=X[1:(n-1)],theta=c(theta1,theta2,theta3),log=TRUE)) }
X <- Eonia[1:85]
X <- na.omit(X)
plot(X)
which(X<0)
is.na(X)
#--- Estima los Parámetros por Máxima Verosimilitud ---#
fit_1=mle(CIR.lik,start=list(theta1=0.1,theta2=0.1,theta3=0.1),
          method="L-BFGS-B",lower=c(0.001,0.001,0.001),upper=c(1,1,1) )

#--- Informe del Ajuste (Estimaciones y Errores Típicos) ---#
summary(fit_1)
# Coefficients:
#           Estimate Std. Error
# theta1 0.1733261 0.06801607
# theta2 0.1934661 0.07861057

```

```
# theta3 0.1618493 0.01379620
```

```
# -2 log L: -96.41743
```

Se debe recordar que este código está diseñado para la parametrización siguiente

$$dr_t = (\theta_1 - \theta_2 r_t)dt + \theta_3 \sqrt{r_t} dW_t,$$

que, en coherencia con la parametrización empleada en este capítulo, habría que llevar a cabo el siguiente cambio

$$\hat{k} = \hat{\theta}_2$$

$$\hat{\theta} = \frac{\hat{\theta}_1}{\hat{k}}$$

$$\hat{\sigma} = \hat{\theta}_3$$

Así, se obtendrían $\hat{k} = 0.1934661$, $\hat{\theta} = 0.8958991$ y $\hat{\sigma} = 0.1618493$.

Y así, si tomamos otra submuestra, *Eo3*, aplicando el código anterior se obtendrían los resultados que siguen

```
### --- Submuestra Eo3 --- ###
```

```
# Coefficients:
```

```
#            Estimate   Std. Error
```

```
# theta1 0.1080756 0.024193232
```

```
# theta2 0.1331019 0.031094402
```

```
# theta3 0.1599054 0.006783222
```

```
# -2 log L: -387.6138
```

Con la reconversión, los estimadores por máxima verosimilitud para *E03* serían $\hat{k} = 0.1331019$, $\hat{\theta} = 0.8119764$ y $\hat{\sigma} = 0.1599054$.

Sin embargo, en las submuestras restantes surgen los problemas ya comentados en la sección anterior.

	khatEo1	thetahatEo1	sigmahatEo1
Máx. Verosim.	0.1934661	0.8958991	0.1618493
Bibby et al	0.1782638	0.8963388	0.1702334
	khatEo3	thetahatEo3	sigmahatEo3
Máx. Verosim.	0.1331019	0.8119764	0.1599054
Bibby et al	0.1377188	0.8122733	0.1684227

Tabla 4.4: Comparación de las estimaciones paramétricas obtenidas para la submuestra $Eo1$ y $Eo3$ mediante el método de aproximación por máxima verosimilitud y la fórmula de Bibby et al (2005).

No obstante, no es éste el único método de estimación; de acuerdo con Bibby, Jacobsen & Sørensen¹⁸ los estimadores de k , θ y σ , suponiendo una muestra de n observaciones de tipos de interés, pueden ser calculados explícitamente mediante las fórmulas:

$$\hat{k}_n = -\ln \frac{(n-1) \sum_{i=2}^n \frac{r_i}{r_{i-1}} - (\sum_{i=2}^n r_i)(\sum_{i=2}^n r_{i-1}^{-1})}{(n-1)^2 - (\sum_{i=2}^n r_i)(\sum_{i=2}^n r_{i-1}^{-1})}, \quad (4.47)$$

$$\hat{\theta}_n = \frac{1}{(n-1)} \sum_{i=2}^n r_i + \frac{e^{-\hat{k}_n}(r_n - r_1)}{(n-1)(1 - e^{-\hat{k}_n})}, \quad (4.48)$$

$$\hat{\sigma}_n^2 = \frac{\sum_{i=2}^n r_{i-1}^{-1} (r_i - r_{i-1} e^{-\hat{k}_n} - \hat{\theta}_n (1 - e^{-\hat{k}_n}))^2}{\sum_{i=2}^n r_{i-1}^{-1}} \left(\left(\frac{\hat{\theta}_n}{2} - r_{i-1} \right) e^{-2\hat{k}_n} - (\hat{\theta}_n - r_{i-1}) e^{-\hat{k}_n} + \frac{\hat{\theta}_n}{2} \right) / \hat{k}_n. \quad (4.49)$$

Bibby, Jacobsen & Sørensen observaron que la existencia de los estimadores arriba formulados existen siempre que $e^{-\hat{k}_n} > 0$ y que ésto sucedía con una probabilidad que tiende a uno cuando $n \rightarrow \infty$.

Con el objeto de efectuar un contraste con los resultados anteriormente arrojados por el método de inferencia por máxima verosimilitud, para las submuestras $Eo1$ y $Eo3$, se observa que las divergencias de los valores son muy leves (ver Tabla 4.5).

Como se percibe en la tabla 4.5, las diferencias en los valores del estimador del parámetro θ son mínimas, siendo éstas en los restantes estimadores, bajas o muy bajas.

¹⁸ «Estimating Functions for Discretely Sampled Diffusion-Type Models», capítulo 4, año 2010.

	Eo1	Eo2	Eo3	Eo4	Eo5	Eo6
ki_hat	0.1782638	1.12751048	0.1377188	0.3452253	0.12716663	0.51771467
thetai_hat	0.8963388	0.09345809	0.8122733	0.3482206	0.09772207	0.08073041
sigmai_hat	0.1702334	0.67590952	0.1684227	0.0332563	0.01534165	0.06975955

	Eo7	Eo8	Eo9	Eo10	Eo11
ki_hat	0.4492095	0.02389449	-0.16548137	0.30577449	0.05435297
thetai_hat	0.1718691	-0.10401890	-0.02798011	-0.12105840	-0.34741672
sigmai_hat	0.1182512	0.18932921	0.05924983	0.04780001	0.01906381

Tabla 4.5: Parámetros estimados de todas las submuestras obtenidas fruto de la partición de la muestra original EONIA (Método de Bibby et al).

Con el fin de obtener los valores de los estimadores de los parámetros para los otros subconjuntos, se recurrirá al último método empleado; así las estimaciones de los diversos parámetros de las correspondientes submuestras se pueden visualizar en la tabla 4.5.

Tras el cálculo de los estimadores de k , θ y σ de cada uno de los grupos obtenidos por la partición de la muestra original EONIA, esto es, \hat{k}_i , $\hat{\theta}_i$ y $\hat{\sigma}_i$, para $i = 1, \dots, 11$; entonces, se podrían aproximar los valores ajustados CIR en el grupo i -ésimo a los tipos de interés trasladados, mediante el esquema Milstein siguiente:

$$\hat{r}_{j+1}^i(k) = \hat{r}_j^i(k) + \hat{k}_i(\hat{\theta}_i - \hat{r}_j^i(k))\Delta + \hat{\sigma}_i\sqrt{\hat{r}_j^i(k)\Delta}Z_{j+1}^i + \frac{1}{2}(\hat{\sigma}_i)^2(\sqrt{\Delta}Z_{j+1}^i)^2 - \Delta, \quad (4.50)$$

```
### ----- CIR fitted values-EONIA's (Milstein) ----- ###
require(sde)
# CIR-Eo1
theta1 = 0.1782638
theta2 = 0.8963388
theta3 = 0.1702334

drift <- expression(theta1 - theta2 * x)
```

```
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(234)
Eo1sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo1sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo2
theta1 = 1.12751048
theta2 = 0.09345809
theta3 = 0.67590952

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(2345)
Eo2sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo2sim, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo3
theta1 = 0.1377188
theta2 = 0.8122733
theta3 = 0.1684227

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(1234)
```

```
Eo3sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo3sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo4
theta1 = 0.3452253
theta2 = 0.3482206
theta3 = 0.0332563

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(11234)
Eo4sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo4sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo5
theta1 = 0.12716663
theta2 = 0.09772207
theta3 = 0.01534165

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(23455)
Eo5sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo5sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")
```

```
# CIR-Eo6
theta1 = 0.51771467
theta2 = 0.08073041
theta3 = 0.06975955

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(123456)
Eo6sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo6sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo7
theta1 = 0.4492095
theta2 = 0.1718691
theta3 = 0.1182512

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(1234567)
Eo7sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo7sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo8
theta1 = 0.02389449
theta2 = -0.10401890
theta3 = 0.18932921
```

```
drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(12345678)
Eo8sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo8sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo9
theta1 = -0.16548137
theta2 = -0.02798011
theta3 = 0.05924983

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(23456789)
Eo9sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                    N = length(x)-1, method = "milstein")
plot(Eo9sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo10
theta1 = 0.30577449
theta2 = -0.12105840
theta3 = 0.04780001

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))
```

```

set.seed(123456789)
Eo10sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                      N = length(x)-1, method = "milstein")
plot(Eo10sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

# CIR-Eo11
theta1 = 0.05435297
theta2 = -0.34741672
theta3 = 0.01906381

drift <- expression(theta1 - theta2 * x)
difussion <- expression(theta3*sqrt(x))
part_deriv <- expression(1/sqrt(x))

set.seed(1122334455)
Eo11sim_adj <- sde.sim(X0=0.1, drift = drift, sigma = difussion, sigma.x = part_deriv,
                      N = length(x)-1, method = "milstein")
plot(Eo11sim_adj, main = "Cox-Ingersoll-Ross (MILSTEIN)")

```

Así, disponemos no sólo de los valores de la serie financiera original, sino que también poseemos los valores ajustados CIR mediante Milstein, y, por ende, con estos ingredientes, estamos en disposición de medir la bondad de ajuste o la precisión del enfoque¹⁹ mediante el error cuadrático medio. Si al ECM, lo denotamos por ϵ^2 , y se define así:

$$\epsilon^2 = \frac{1}{n} \sum_{i=1}^n e_i^2.$$

Donde $e_i = r_i - \hat{r}_i$, siendo r_i , los tipos de interés de la muestra original del grupo i -ésimo, y, \hat{r}_i , los respectivos valores ajustados obtenidos mediante el esquema de aproximación de Milstein 4.50. Así, pues la diferencia entre ambos, es decir, e_i , denota el residuo del i -ésimo grupo.

¹⁹Asimismo, también podríamos obtener cualquier otra medida de la bondad de ajuste, i.e. R^2 .

No deseo finalizar esta sección sin mencionar que existen otros métodos de obtención de los estimadores de los parámetros. Así, quiero destacar uno que emplea los modelos ARIMA para la obtención de los residuos estandarizados²⁰ que deberían seguir una distribución normal estándar. Posteriormente, tras chequear efectivamente que estos residuos estandarizados están normalmente distribuidos, se emplearían como valores de los Z_{j+1}^i usados en la aproximación de Milstein. Por ello, aunque sea un proceso tangencial, no nuclear, del tema en desarrollo, me gustaría dedicarle unas líneas para una breve exposición del mismo.

Esta metodología, que podría ahorrar costes computacionales según los creadores del modelo CIR#, consiste en usar los residuos normales estandarizados de un modelo ARIMA óptimo (en términos de BIC o/y AIC) respecto a los datos muestrales, para estimar los parámetros k_i , θ_i y σ_i , de los respectivos sub-grupos obtenidos con la partición, y para la aproximación por Milstein de los valores del modelo CIR ajustado por los estimadores de los parámetros previamente obtenidos.

Por ello, para derivar el modelo ARIMA(p, d, q) que mejor ajusta los datos de los diferentes conjuntos resultantes de la partición de la muestra original, EONIA, se emplea la función de R, *auto.arima()* aplicada a cada uno de los conjuntos Eo_i , para $i : 1, \dots, 11$, seleccionamos un modelo ARIMA(p, d, q) cuyos residuos estandarizados satisfacen una serie de condiciones²¹.

Para el coeficiente de difusión $\sigma\sqrt{r_t}$, donde σ , la volatilidad, ésta se estima mediante $\hat{\sigma}_i$, que es la desviación típica del grupo i -ésimo. Para el parámetro drift o $k(\theta - r_t)$, se necesita estimar θ o la media a largo plazo, mediante $\hat{\theta}_i$, la media del grupo i -ésimo; así, los resultados serían

El parámetro k o la velocidad de reversión a la media, se estima mediante \hat{k}_i que es el valor resultante de resolver el siguiente problema de optimización no-lineal y no-convexo

$$\min_{k>0} S_i(k),$$

donde

²⁰El proceso de Wiener, en la parte estocástica o aleatoria del proceso CIR, es reemplazado por los residuos estandarizados normalmente distribuidos de los modelos óptimos ARIMA(p, d, q) apropiadamente elegidos.

²¹Las condiciones a satisfacer serían: la ausencia tanto de autocorrelación (AC) como de autocorrelación parcial (PAC) en la serie financiera, EONIA; la serie tiene que ser estacionaria y los residuos estandarizados tienen que estar normalmente distribuidos.

	Eo1	Eo2	Eo3	Eo4	Eo5	Eo6
thetai_hat	0.8960471	0.3718762	0.8203135	0.3538696	0.1016353	0.08039015
sigmai_hat	0.2470699	0.07613589	0.2716667	0.03707535	0.01106673	0.01884667

	Eo7	Eo8	Eo9	Eo10	Eo11
thetai_hat	0.1749146	0.130243	-0.02508784	-0.1195537	-0.3435543
sigmai_hat	0.05754809	0.1345367	0.04095691	0.02207345	0.03572262

Tabla 4.6: Los valores de los parámetros volatilidad, σ , y media a largo plazo, θ , estimados mediante un simple cálculo de la desviación típica y de la media aritmética, respectivamente.

$$S_i(k) := \sqrt{\frac{\sum_{j=1}^{n_i} (u_j^i(k) - \bar{u}^i(k))^2}{n_i - 1}}, \quad \text{para cualquier } k > 0, \quad j = 1, \dots, n_i.$$

Donde

$$u_j^i(k) := r_j^i(k) - r_{trasl,j}^i,$$

con $r_j^i(k)$ siendo una muestra de valores simulados de los tipos de interés de un proceso CIR como función del parámetro teórico de la velocidad de reversión, k , obtenida mediante una aproximación de un esquema de Milstein ; y $r_{trasl,j}^i$, la muestra de EONIA trasladada mediante la adición de α , esto es, $r_{trasl}(t) = r_{eonia}(t) + 0.366$, con $t \in [0, T]$.

La serie financiera simulada por aproximación mediante Milstein simplemente hace uso del lema de Itô para incrementar la precisión de la aproximación al añadir el término de segundo orden; así, como ya sabemos, la aproximación de Milstein sería

$$Y_{t+1} = Y_t + b(t_i, Y_i)\Delta t + \sigma(t_i, Y_i)\Delta W_t + \sigma_x(t_i, Y_i). \quad (4.51)$$

Donde $\sigma_x(t_i, Y_i)$ es la derivada parcial de $\sigma(x, t)$ respecto de x .

Así, reparametrizando la ecuación , tendríamos:

$$r_{j+1}^i(k) = r_j^i(k) + k(\hat{\theta}_i - r_j^i(k))\Delta + \hat{\sigma}_i \sqrt{r_j^i(k)\Delta} Z_{j+1}^i + \frac{1}{2}(\hat{\sigma}_i)^2 (\sqrt{\Delta} Z_{j+1}^i)^2 - \Delta,$$

donde $\Delta = 1$ (tipos de interés diarios) y Z_{j+1}^i son las v.a. normales estándar. En este punto es donde reside la novedad de esta metodología, pues, los autores usan los residuos estandarizados normales de los modelos ARIMA(p, d, q) óptimos como los Z_{j+1}^i , no sólo para la aproximación por Milstein para la obtención de \hat{k} , sino que también, una vez estimado el parámetro de reversión a la media, k , por un algoritmo de optimización, y junto con los parámetros previamente estimados de θ y σ , se procedería análogamente al estudio de la bondad de ajuste, ϵ^2 .

Por tanto, ya que disponemos de unos estimadores de los parámetros drift y de difusión para cada submuestra, ¿cómo podemos llegar a los estimadores de la muestra original, EONIA?

En principio, éstos se podrían estimar mediante una sencilla ponderación, es decir, otorgando unos pesos a cada uno de los estimadores submuestrales, en función del tamaño de cada submuestra respecto a la muestra completa. Así,

$$\hat{k}_{CIR} = \sum_{i=1}^{11} w_i \hat{k}_i,$$

$$\hat{\theta}_{CIR} = \sum_{i=1}^{11} w_i \hat{\theta}_i$$

y

$$\hat{\sigma}_{CIR} = \sum_{i=1}^{11} w_i \hat{\sigma}_i.$$

Con $w_i = \frac{n_i}{n}$; donde n_i es el tamaño de la i -ésima submuestra, y n , el tamaño de la muestra EONIA.

Tras el cálculo de los pesos, w_i , los estimadores de los parámetros de la muestra completa son

$$\hat{k}_{CIR} = 0.2872801$$

$$\hat{\theta}_{CIR} = 0.05324872$$

$$\hat{\sigma}_{CIR} = 0.1466413$$

La solución CIR única a la ecuación diferencial estocástica $dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t$ es

$$r_t = (r_0 - \hat{\theta}_{CIR})e^{-\hat{k}_{CIR}t} + \hat{\sigma}_{CIR}e^{-\hat{k}_{CIR}t} \int_0^t e^{\hat{k}_{CIR}u} \sqrt{r_u} dW_u.$$

Por ende, la solución CIR única a la EDE anterior en nuestro caso es

$$r_t = (r_0 - 0.05324872)e^{-0.2872801t} + 0.1466413e^{-0.2872801t} \int_0^t e^{0.2872801u} \sqrt{r_u} dW_u.$$

Para el caso concreto de $r_0 = 0.1013$ y $t = 2549$ se tiene que la solución explícita vendría dada por la expresión que sigue

$$r_{2549} = 0.04805128e^{-0.2872801t} + 0.1466413e^{-0.2872801t} \int_0^{2549} e^{0.2872801u} \sqrt{r_u} dW_u.$$

4.2. Conclusión

Como ya hemos visto en capítulos anteriores, en finanzas, un modelo en tiempo continuo se basa en uno o más procesos de difusión estacionarios $\{X_t, t \geq 0\}$, con unas dinámicas representadas por las ecuaciones diferenciales estocásticas de Itô de la forma:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad (4.52)$$

donde $\{W_t, t \geq 0\}$ es un movimiento Browniano estándar, $\mu(\cdot)$ y $\sigma(\cdot)$ son funciones drift y difusión del proceso, respectivamente.

Generalmente, éstas se parametrizan como sigue:

$$\mu(x) = \mu(x, \theta) \quad y \quad \sigma^2(x, \theta), \quad \text{donde } \theta \in \Theta \subset \mathbb{R}^k.$$

Asimismo, ya hemos visto que la elección de las familias de los parámetros de drift y difusión varían según el modelo.

Dado que perseguimos mostrar la aplicabilidad de las ecuaciones diferenciales estocásticas a las finanzas, nos decantamos por analizar cómo se comportan ante una variable económico-financiera.

Así, con el objetivo de intentar modelar, a partir de una serie financiera real, la que es quizás la variable económico-financiera más importante, esto es, el tipo de interés, sin que se altere la naturaleza estocástica del proceso que describe la dinámica de los tipos de interés, se ha empleado el modelo de Cox-Ingersoll-Ross. El proceso CIR, desde su aparición, gozó de una enorme popularidad entre la comunidad científica pues es un proceso relativamente sencillo de manejar, es un modelo unifactorial y de reversión a la media, y, adicionalmente, nos provee con una solución explícita siempre que se cumplan determinadas restricciones paramétricas, a recordar $2\theta_1 > \theta_2^2$, θ_1, θ_2 y $\theta_3 \in \mathbb{R}_+$; lo cual es absolutamente maravilloso. Además, supuso una presunta mejora sobre el modelo de Ornstein-Uhlenbeck el cual inadmisiblemente podía alcanzar valores negativos. Sin embargo, mediante el uso de una serie financiera como la empleada por nosotros, con valores de los tipos de interés diarios durante un período que abarca los años posiblemente más críticos en las finanzas mundiales, los resultados de inferencia paramétrica obtenidos no hemos comprobado si son o no aceptables, pues no era éste el objetivo del presente trabajo, las dificultades que nos hemos encontrado para aplicar la diversa metodología expuesta en capítulos anteriores, han sido mayúsculas.

De hecho, hemos visto que, si se usa la serie financiera EONIA completa, el método de estimación de los parámetros mediante verosimilitud exacta fracasa. Una de las principales razones estriba en el hecho de que la densidad transitoria o condicional del proceso CIR sigue una distribución Chi-cuadrada no-centrada lo que implícitamente conlleva en primer lugar, una restricción en los valores de las observaciones, éstos necesariamente tienen que ser al menos no-negativos, y, en segundo lugar, existe un riesgo computacional de explosión tanto si se utiliza la función Bessel reescalada exponencialmente como si se emplea la densidad Chi-cuadrada. Asimismo, hemos mostrado que si se usan determinados fragmentos de la muestra original, esta metodología devuelve los valores de los estimadores con un error estándar bajo. Si nos preguntamos la causa de que una misma metodología sea o no efectiva, nuestro estudio revela que en aquellas submuestras donde el método de inferencia de la verosimilitud exacta funciona, precisamente es en aquellas que contienen valores estrictamente positivos.

Con el objeto de solventar estos desafíos, una de las primeras medidas que se ha implementado consiste en transformar la serie original mediante una traslación. El valor de esta constante se logra a través de un método recomendado por Orlando, Mininni y Bufalo. Una vez que se ha trasladado la serie y comprobado que efectivamente todos los valores son positivos, ya podemos hacer uso de las diferentes metodologías para la obtención de los estimadores

de los parámetros teóricos.

Otro desafío importante era la propia muestra. El modelar una muestra de 2548 valores observados durante un período de diez años convulsos en las finanzas europeas invitaba al pesimismo. Parece razonable pensar que la volatilidad al principio del período, con unos tipos de interés más altos, no puede ser la misma que al final del período, con unos tipos de interés muy bajos. Dadas estas posibles divergencias en la volatilidad, parecería lógico llevar a cabo una partición de la muestra original y proceder a un análisis en cada submuestra y observar cómo se comporta el proceso CIR en cada una de ellas.

Una vez que se ha particionado la muestra en 11 subconjuntos, se han aplicado diferentes métodos de inferencia paramétrica con los consiguientes resultados ya reflejados con anterioridad. Comentar que hemos vuelto a encontrar dificultades con el método de inferencia por verosimilitud exacta e ignoramos la génesis exacta de estos problemas en este caso particular.

Finalmente, dado que las submuestras poseen tamaños muy diferentes, hemos decidido aportarles unos pesos a los valores de los estimadores basados en una simple ponderación respecto al tamaño de la muestra original EONIA. Con estos valores, estamos en disposición de obtener una solución explícita CIR.

Por otro lado, también nos planteamos analizar el comportamiento de un modelo más sencillo como el proceso de Ornstein-Uhlenbeck ante una muestra como la de EONIA. Debo anticipar que los resultados arrojados por el modelo de O-U han sido menos laboriosos de consecución. Con el ánimo de no incurrir en reiteraciones innecesarias, los métodos de inferencia paramétrica funcionan debidamente. Así, el método de inferencia de verosimilitud exacta arroja unos resultados parecidos a los logrados por el 'método' de Bibby y otros; concretamente, los valores de los estimadores del parámetro θ_1 y θ_2 son similares; la disparidad nace con el estimador de θ_3 que representa la volatilidad. Esto posiblemente se deba a una de las razones ya expuestas: la variabilidad de la volatilidad en una serie financiera tan grande durante un período tan convulso desde el punto de vista financiero. No olvidemos que θ_3 en el proceso de O-U es constante y no depende del espacio de estados.

Durante el estudio del modelo de Ornstein-Uhlenbeck no hemos querido realizar la fragmentación de la serie original EONIA por cuestiones didácticas, ya que, en el capítulo siguiente al del proceso O-U, se iba a efectuar este método y no reside en la naturaleza de este trabajo la estimación de los parámetros teóricos, sino que el fin último del mismo es demostrar que las ecuaciones diferenciales estocásticas suponen una herramienta muy potente

para modelar variables económico-financieras como los tipos de interés, quedando así probado desde un punto de vista práctico la enorme aplicabilidad que las EDE's poseen en un ámbito económico-financiero.

Bibliografía

- [1] Steel J. M. *Stochastic Calculus and Financial Applications, Series Applications of Mathematics*. Springer-Verlag, 2001.
- [2] Clebaner F. C. *Introduction to Stochastic Processes with Applications*. Imperial College Press, 1999.
- [3] Rebus D. and M. Yor. *Continuous Martingales and Stochastic Calculus*. Springer, 1991.
- [4] Vasicek O. *An Equilibrium Characterization of the Term Structure*. Journal of Financial Economics, 1977.
- [5] Cox J. C., Ingersoll J. E., and Ross S. A. *A Theory of the Term Structure of Interest Rates*. Econometrica, 1985.
- [6] Elliot R. J. and Kopp P. E. *Mathematics of Financial Markets*. Springer-Verlag, 1999.
- [7] Merton R. C. *Theory of Rational Option Pricing*. Bell Journal of Economic and Management Science, 1973.
- [8] Black F. and Scholes M. *The Pricing of Options and Corporate Liabilities*. The Journal of Political Economy, 1973.
- [9] Hull J. C. *Options, Futures and other Derivatives*. Pearson, 2012.
- [10] Manuel Moreno, Pedro Serrano, and Winfried Stute. Statistical properties and economic implications of jump-diffusion processes with shot-noise effects. *European Journal of Operational Research*, 214(3):656–664, 2011.
- [11] Allen L., Boudoukh J., and Saunders A. *Quantifying Volatility in VaR Models*. Blackwell Publishing Ltd., 2004.
- [12] Chan K. C., Karolyi G. A., Longstaff F. A., and Sanders A. B. An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance*, 47(3):1209–1227, 1992.

-
- [13] Iacus S. M. *Simulation and Inference for Stochastic Differential Equations*. Springer, 2008.
- [14] Zipf R. *Fixed Income Mathematics*. Academic Press, 2003.
- [15] Ait-Sahalia Yacine. Testing continuous-time models of the spot interest rate. *The Review of Financial Studies*, 9(2):385–426, 1996.
- [16] Mwamba J. M., Thabo L., and Uwilingiye J. Modelling the short-term interest rate with stochastic differential equation in continuous time: Linear and nonlinear models. *MPRA*, (64386):385–426, 2015.
- [17] Orlando G., Mininni R. M., and Bufalo M. On the calibration of short-term interest rates through a cir model. 2018.
- [18] Orlando G., Mininni R. M., and Bufalo M. Forecasting interest rates through vasicek and cir models: A partitioning approach. 2019.
- [19] Venegas Martínez F. *Riesgos Financieros and Económicos*. CENGAGE Learning, 2008.
- [20] Fabozzi F. J. *Fixed Income Mathematics: Analytical and Statistical Techniques*. McGraw-Hill, 2006.
- [21] Elerian O. A note of the existence of a closed form conditional density for the milstein scheme. *Working Paper*, 1998.
- [22] Shoji L. and Ozaki T. Estimation for nonlinear stochastic differential equations by a local linearization method. *Stochastic Anal. Applications*, 16:733–752, 1998.
- [23] Lavielle M. Using penalized constrasts for the change-point problem. *Signal Processing*, 2005.
- [24] Lipster R. S. and Shiriyayev A. N. *Statistics of Random Processes*. Springer-Verlag, 1997.