

Entropía relativa y riesgo de modelo en Swaps de tipos de interés



Cristina Arribas Francisco

Trabajo de fin de máster en Ingeniería Matemática
Universidad Complutense de Madrid

Director del trabajo: Daniel Arrieta Rodríguez
Tutor académico: Gerardo Enrique Oleaga Apadula
Septiembre, 2018

Índice general

	Página
Agradecimientos	I
Resumen/Abstract	II
1. Introducción y planteamiento del problema	1
1.1. Introducción: La aparición del smile de volatilidad	1
1.2. Planteamiento del problema	4
2. El modelo de Vasicek para el tipo a corto plazo	6
2.1. La formulación del tipo a corto plazo, $r(t)$	6
2.2. Una expresión para los precios de los bonos, $P(t, T)$	11
3. La metodología de Entropía Relativa Mínima (ERM) y Weighted Monte Carlo (WMC)	15
3.1. El concepto de entropía relativa	15
3.2. Entropía Relativa Mínima y simulación Monte Carlo: Monte Carlo Ponderado	18
3.2.1. Idea general de la estrategia a seguir	18
3.2.2. Desarrollo teórico	21
4. El riesgo de modelo cuantificado mediante métricas de exposición	32
5. Valoración de un swap de tipos de interés con <i>smile</i> y análisis del riesgo de modelo	34
5.1. Formulación del problema a resolver	34
5.2. Simulación del tipo de interés a corto plazo mediante el modelo de Vasicek .	35
5.3. Obtención de una medida de probabilidad con <i>smile</i>	36
5.3.1. Resultados obtenidos	40
5.4. Valoración del swap	41
5.5. Análisis del riesgo de modelo mediante métricas de exposición	44
6. Conclusiones y líneas futuras de investigación	49
6.1. Conclusiones	49
6.2. Futuras líneas de investigación	50
Anexo A: Teoremas Fundamentales de Valoración Financiera	51
Bibliografía	52

Agradecimientos

A Daniel Arrieta y a Gerardo Oleaga, por toda la dedicación y ayuda que me han prestado en la realización de este trabajo. A todos los profesores de la Facultad de Ciencias Matemáticas de la Universidad Complutense de Madrid, de los que tanto he aprendido durante los años del Grado y del Máster. A mi familia y amigos, por el apoyo incondicional desde siempre.

Resumen

Las volatilidades que cotizan en el mercado para opciones *out the money* corresponden a precios más altos que los proporcionados por la fórmula de Black-Scholes. Esta característica se conoce como *smile* de volatilidad y conviene considerarla en los modelos de valoración de contratos financieros. Proponemos una metodología para instrumentos derivados de los tipos de interés que permite obtener una probabilidad neutral al riesgo incorporando de manera implícita la influencia del *smile* de volatilidad. Así, esta medida posibilita la valoración correcta de derivados mediante el Primer Teorema Fundamental de Valoración Financiera. En concreto, en este Trabajo la empleamos para valorar un swap. Posteriormente se lleva a cabo un análisis del riesgo de modelo asociado a la no consideración del *smile* de volatilidad.

Abstract

Market volatilities for out the money options correspond to higher market prices than those given by the Black-Scholes formula. This fact is known as volatility smile and should be considered in mathematical models for pricing financial instruments. We propose a methodology for interest rate derivatives which determines a risk-neutral measure that implicitly includes the influence of volatility smile. Thus, this probability can be used, along with the First Fundamental Theorem of Asset Pricing, to give a correct price to derivatives. In particular, in this project we price a swap. Finally, an analysis is performed to determine model risk when volatility smile is not considered.

Capítulo 1

Introducción y planteamiento del problema

1.1. Introducción: La aparición del smile de volatilidad

En 1973 Fischer Black y Myron Scholes publicaron un modelo para la valoración de opciones europeas bajo la hipótesis de que el precio del subyacente, S , evoluciona según una dinámica Lognormal¹. Este modelo ha sido ampliamente aceptado y utilizado en los mercados financieros desde aquel momento. Sin embargo, el lunes 19 de octubre de 1987 (conocido como *lunes negro*) las Bolsas de todo el mundo sufrieron un gran desplome (por ejemplo, el índice Dow Jones Industrial Average cayó en 508 puntos básicos, es decir, un 22.6%). La ocurrencia de este evento extremo (con bajísima probabilidad de ocurrir) propició la necesidad de llevar a cabo correcciones a los precios que el modelo de Black-Scholes daba a las opciones, de manera que se tuvieran más en cuenta escenarios como este. La idea consistía en subir ligeramente los precios de las opciones cuyos strikes estaban muy alejados del precio actual del subyacente (opciones *Out The Money*², que denotaremos por *OTM* de ahora en adelante), pues son las que se podrían utilizar para la cobertura de eventos extremos que tuvieran lugar. La manera en que dichas correcciones se empezaron a realizar derivó de los parámetros que aparecían en la expresión del modelo de Black-Scholes. Para entenderlo mejor, recordemos la expresión propuesta en este modelo para valorar una opción europea cuando el subyacente no paga dividendos.

Notación. Denotaremos a los parámetros que intervienen en el modelo de Black-Scholes de la siguiente manera:

- T : Tiempo de vencimiento de la opción.
- K : Strike o precio de ejercicio de la opción.
- B_T : Valor a tiempo T de una unidad monetaria depositada en la cuenta bancaria en el instante 0. Si r es el tipo de interés libre de riesgo y considerando composición continua, tenemos que $B_T = e^{rT}$.

¹En realidad, esta hipótesis no se cumple exactamente en el mercado. Por ello, al usar este modelo existe un riesgo asociado a la violación de las hipótesis del mismo, el cual es un tipo de riesgo de modelo.

²Es frecuente la terminología en inglés en esta área, por lo que algunos conceptos aparecerán con su denominación original.

- S_0 : Valor del activo subyacente en el instante inicial.
- σ : Volatilidad del subyacente.
- ψ : Parámetro que toma el valor 1 si la opción que se quiere valorar es una call (de compra) o el valor -1 si es una put (de venta).

Entonces, la fórmula de Black-Scholes para obtener el valor a tiempo 0, V_0 , de opciones europeas call o put sobre un subyacente que no paga dividendos es la siguiente:

$$V_0 = e^{-rT} \cdot \psi \cdot [B_T S_0 \cdot \Phi(\psi \cdot d_1) - K \cdot \Phi(\psi \cdot d_2)],$$

donde:

$$d_1 := \frac{\log(B_T S_0 / K)}{\sigma \sqrt{T}} + \frac{\sigma \sqrt{T}}{2}, \quad d_2 := \frac{\log(B_T S_0 / K)}{\sigma \sqrt{T}} - \frac{\sigma \sqrt{T}}{2}$$

y Φ es la densidad acumulada de la distribución Normal estándar.

Como podemos observar, el único parámetro libre en la expresión de V_0 es σ , la volatilidad del subyacente, que se puede estimar a partir de datos históricos. También se puede obtener la *volatilidad implícita*, que es el valor de σ para el que la fórmula de Black-Scholes nos da el precio de mercado exacto. Si el modelo de Black-Scholes fuera totalmente cierto, la volatilidad histórica sería muy similar a la implícita y tomaría el mismo valor para todos los strikes K . Intuitivamente, si aumenta la volatilidad del activo subyacente, también lo harán los precios de las opciones sobre ese activo, incluidos los de las opciones OTM. Esto es así porque es más probable que se den eventos extremos y, por tanto, que se tengan que utilizar estas opciones para cubrirlos. Las correcciones en los precios de las opciones, entonces, se llevarán a cabo aumentando la volatilidad de las opciones OTM, es decir, la volatilidad asociada a strikes alejados del valor actual del subyacente. Para ilustrar esto, veamos un ejemplo sencillo con una opción put.

Ejemplo 1. Las opciones put protegen al poseedor de bajadas en los precios del activo subyacente, ya que le dan derecho a vender este por un precio fijo, el strike K . Por ello, las put con strikes muy bajos estarán asociadas a mayores valores de la volatilidad del subyacente, que son los que permitirán que se dé un evento extremo en el precio de este y, por tanto, ejercerlas en un vencimiento determinado. Gráficamente, tendríamos lo siguiente:

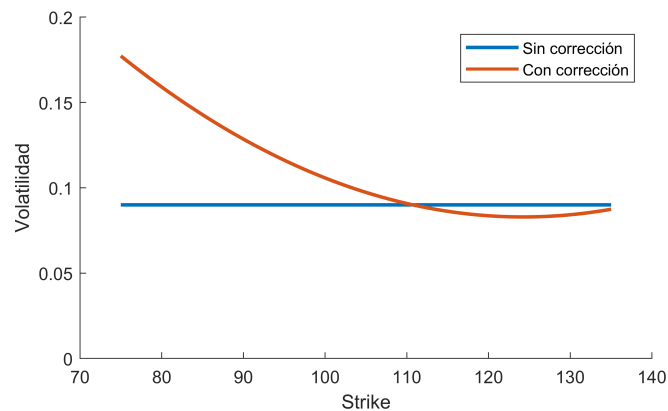


Figura 1.1: Volatilidades en función de los strikes

En color azul se representa la situación en la que los precios de las opciones OTM no son corregidos. Simplemente, se ha estimado la volatilidad del subyacente (en este caso, $\sigma = 0.09$) y se usa en la ecuación de Black-Scholes como un dato constante para todos los strikes. Sin embargo, en color naranja observamos la situación contraria: cada strike lleva asociada una determinada volatilidad, que será la que se emplee en la expresión de Black-Scholes para determinar el precio de la opción.

La curva que hemos obtenido en la figura 1.1 se denomina **smile de volatilidad**, debido a su forma. Cabe destacar que cada smile de volatilidad está asociado a un determinado vencimiento, T , de las opciones. Si considerásemos los smiles de volatilidad para varios vencimientos distintos, obtendríamos lo que se conoce como **superficie de volatilidad**.

Actualmente en el mercado cotizan activos cuyos precios ya están corregidos e incorporan el smile de volatilidad. Por ello, se hace necesario considerar esta característica en los modelos que se emplean en la industria financiera. De no hacerlo, los resultados y conclusiones obtenidas con ellos pueden no ser válidos, ya que infravalorarán la ocurrencia de eventos extremos (como el citado *lunes negro*).

Para tener en cuenta la influencia del smile en los modelos se debe dar más peso a la ocurrencia de dichos eventos extremos, de manera que se modifique ligeramente la forma de las distribuciones que intervienen, haciéndolas más leptocúrticas (más apuntadas en la zona central y con colas más pesadas). La figura 1.2, donde se muestra la cola izquierda de un posible histograma de rentabilidades, muestra cómo una distribución leptocúrtica recoge mejor los valores más extremos que una distribución mesocúrtica (como la Normal).

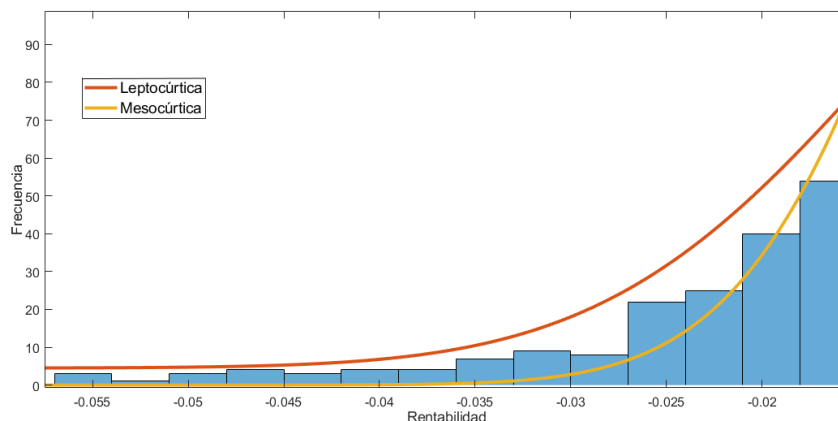


Figura 1.2: Comparativa entre distribuciones leptocúrtica y mesocúrtica

A lo largo de los años se han presentado diversas soluciones al problema de la incorporación del smile de volatilidad en los modelos financieros. Usualmente, se ha tratado de mejorar la modelización de la evolución del subyacente que aparece en el modelo de Black-Scholes añadiendo algún término adicional o generalizando los existentes. Podemos citar el modelo de Volatilidad Estocástica (Hull-White), el modelo de Volatilidad Local (Dupire) y el modelo de Saltos (Merton). A modo de ejemplo, recordemos que la ecuación que apa-

rece en el modelo de Black-Scholes es $\frac{dS_t}{S_t} = rdt + \sigma dW_t$ (donde dW_t es un movimiento Browniano), mientras que el modelo elaborado por Dupire introduce una función, $\sigma(S_t, t)$, para la volatilidad. Sin embargo, en este Trabajo vamos a enfocar el problema desde una metodología no paramétrica, como se comentará en la siguiente sección.

1.2. Planteamiento del problema

Los **swaps de tipos de interés** son instrumentos financieros muy extendidos. Estos contratos involucran a dos partes y se emplean para intercambiar intereses (en muchas ocasiones, para pasar de interés variable a fijo y viceversa). La mecánica de los swaps es muy sencilla: se fijan unas fechas de pago y, en ellas, cada parte pagará a la otra parte un interés (fijo o variable, según se haya fijado con anterioridad al comienzo del swap) sobre un mismo nominal. El nominal no se intercambia, solo se pagan los intereses. Por lo anterior, el subyacente en los swaps es un tipo de interés.

La valoración de los swaps, como la de cualquier otro instrumento, se ve afectada por la problemática del **smile de volatilidad** que hemos explicado en la sección anterior. Como hemos comentado, para los casos en los que el subyacente es de *equity*, se han desarrollado modelos que pretenden dar alguna solución a la valoración con smile. También para esos casos, existe un enfoque no paramétrico que se conoce como metodología de Entropía Relativa Mínima (ERM), que ha demostrado ofrecer muy buenos resultados. El hecho de que esta técnica sea realmente eficaz para los casos en que el subyacente es de *equity* ha motivado la realización de este Trabajo. A lo largo de él, vamos a determinar si la aplicación de ERM al caso en que el subyacente es un tipo de interés también es factible y brinda resultados positivos. Para este caso particular no hemos identificado investigación previa.

El recorrido que se ha planteado para realizar este trabajo de investigación se puede dividir en dos fases principales. La primera consiste en obtener unas probabilidades que incorporen de manera implícita el smile de volatilidad presente en el mercado. En esta fase se aplicará la técnica de *Weighted Monte Carlo* (WMC), que es una forma de poner en práctica la metodología de Entropía Relativa Mínima. Estas probabilidades se usarán en la segunda fase para valorar el swap. Se simularán precios de este instrumento normalmente y, a la hora de obtener el resultado final, se ponderarán los resultados de cada simulación con esas probabilidades. Si ponderásemos de manera equiprobable (aplicando la simulación Monte Carlo habitual), no estaríamos teniendo en cuenta la problemática del smile de volatilidad. Por ello, en esta fase es donde se realiza el análisis del riesgo de modelo estudiando los beneficios derivados de incorporar el smile de volatilidad en la valoración del swap. Esto lo haremos mediante la evaluación de algunas métricas de exposición en los dos casos que acabamos de comentar: cuando se hace la simulación Monte Carlo habitual y cuando esta se lleva a cabo aplicando las probabilidades que incorporan el smile (Weighted Monte Carlo).

El objetivo final que se persigue con este Trabajo de Fin de Máster es, entonces, obtener y describir una metodología que permita valorar instrumentos financieros (dependientes de los tipos de interés) incorporando la presencia del smile de volatilidad. Para ello, nos basaremos en la metodología de Entropía Relativa Mínima, aplicándola mediante simulación Monte

Carlo Ponderada. Igualmente, se pretende concretar la influencia del smile en la valoración y la importancia de la introducción del mismo en los modelos financieros para minimizar el riesgo de modelo (el que se deriva del uso de un modelo inadecuado). En particular, todo lo anterior se pondrá en práctica en el caso concreto de la valoración de un swap de tipos de interés.

Capítulo 2

El modelo de Vasicek para el tipo a corto plazo

Como hemos comentado en el capítulo introductorio, en este Trabajo vamos a aplicar la metodología de Entropía Relativa Mínima a instrumentos dependientes de los tipos de interés. Por este motivo, se dedica este capítulo al estudio de un modelo que permite simular el tipo de interés a corto plazo: el modelo de Vasicek.

2.1. La formulación del tipo a corto plazo, $r(t)$

La **estructura temporal de los tipos de interés** (ETTI) es una curva que representa de manera gráfica los tipos de interés existentes en el mercado a tiempo t para diferentes plazos de longitud τ . Es decir, la ETTI vista a tiempo t es la función $\tau \rightarrow R_t(\tau)$, donde $R_t(\tau)$ representa el tipo de interés de mercado de plazo τ que hay a tiempo t (para aplicar en el intervalo $[t, t + \tau]$). Cuando τ es un plazo arbitrariamente corto, entonces el tipo de interés se denomina **tipo a corto plazo** o, con la denominación inglesa, *short rate*. Si denotamos por $r(t)$ a este tipo a corto plazo, se tiene que $r(t) = \lim_{\tau \rightarrow 0} R_t(\tau)$. Por tanto, $r(t)$ es el tipo de interés (instantáneo) que se observa en el mercado a tiempo t para un plazo temporal muy pequeño.

El **modelo de Vasicek** es un proceso de reversión a la media¹ (o proceso de Ornstein-Uhlenbeck) que propone una expresión para la evolución del tipo de interés a corto plazo.

Sea $W(t)$, $t \geq 0$, un movimiento Browniano. Entonces, el modelo de Vasicek introduce la siguiente ecuación diferencial estocástica para la evolución del tipo de interés a corto plazo, $r(t)$:

$$dr(t) = \kappa(\theta - r(t))dt + \sigma dW(t), \quad (2.1)$$

donde θ es el nivel de reversión a la media, κ es la velocidad de reversión a la media y σ es

¹Intuitivamente, un proceso de reversión a la media es aquel que, a lo largo del tiempo, tiende a moverse hacia un determinado nivel fijo. Es decir, si se ha producido alguna subida por encima de dicho nivel, tenderá a bajar hasta él a medida que pase el tiempo (y viceversa). Este movimiento de reversión a la media puede ocurrir a diferentes velocidades.

la volatilidad de la short rate² (los tres son constantes en este modelo). El modelo parte de la condición inicial $r(0) = r_0$.

Una de las principales ventajas del modelo de Vasicek es que nos permite determinar una expresión cerrada para la solución, $r(t)$. Veamos cómo llegar a ella, basándonos en la estrategia propuesta en [11]. Para ello, necesitaremos los siguientes conceptos teóricos:

Definición 1 (Proceso de Itô). *Dado $W(t)$, con $t \geq 0$, decimos que un proceso estocástico $X(t)$ es un proceso de Itô cuando es de la forma:*

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)d(u), \quad (2.2)$$

donde $X(0)$ es constante, $\Delta(t)$ y $\Theta(t)$ son procesos estocásticos adaptados y las integrales involucradas están bien definidas y son finitas.

Observación. *Si escribimos lo anterior con notación diferencial, el proceso de Itô $X(t)$ tiene la forma:*

$$dX(t) = \Delta(t)dW(t) + \Theta(t)d(t). \quad (2.3)$$

Para obtener la expresión cerrada de $r(t)$ nos ayudaremos de la fórmula de Itô-Doeblin:

Teorema 1 (Fórmula de Itô-Doeblin). *Sea $X(t)$, con $t \geq 0$, un proceso de Itô y sea $f(t, x)$ una función para la que existen $f_t(t, x)$, $f_x(t, x)$ y $f_{xx}(t, x)$ y son continuas. Entonces, para todo $t \geq 0$, se tiene que:*

$$df(t, X(t)) = f_t(t, X(t))dt + f_x(t, X(t))dX(t) + \frac{1}{2}f_{xx}(t, X(t))dX(t)dX(t). \quad (2.4)$$

Proposición 1. *La solución de la ecuación del modelo de Vasicek es:*

$$r(t) = e^{-\kappa t}r(0) + \int_0^t e^{\kappa(s-t)}\kappa\theta ds + \int_0^t e^{\kappa(s-t)}\sigma dW(s). \quad (2.5)$$

Demostración. Tenemos que $r(t)$ es un caso particular de proceso de Itô, ya que es de la forma de (2.3). Entonces, vamos a aplicar la fórmula de Itô-Doeblin con la función $f(t, x) = e^{\kappa t}x$. Se tiene que:

$$\begin{aligned} df(t, r(t)) &= d(e^{\kappa t}r(t)) = \kappa e^{\kappa t}r(t)dt + e^{\kappa t}dr(t) \Leftrightarrow \\ &\Leftrightarrow d(e^{\kappa t}r(t)) = \kappa e^{\kappa t}r(t)dt + e^{\kappa t}(\kappa(\theta - r(t))dt + \sigma dW(t)) \Leftrightarrow \\ &\Leftrightarrow d(e^{\kappa t}r(t)) = \kappa e^{\kappa t}r(t)dt + e^{\kappa t}\kappa\theta dt - e^{\kappa t}\kappa r(t)dt + e^{\kappa t}\sigma dW(t) \Leftrightarrow \\ &\Leftrightarrow d(e^{\kappa t}r(t)) = e^{\kappa t}\kappa\theta dt + e^{\kappa t}\sigma dW(t). \end{aligned}$$

Integrando entre 0 y t y operando, llegamos a la expresión cerrada para $r(t)$:

$$\int_0^t d(e^{\kappa s}r(s)) = \int_0^t e^{\kappa s}\kappa\theta ds + \int_0^t e^{\kappa s}\sigma dW(s) \Leftrightarrow$$

²Para ser precisos, es una estimación de la desviación típica de las variaciones del tipo a corto plazo (ecuación (2.1)).

$$\begin{aligned}
&\Leftrightarrow e^{\kappa t}r(t) - e^{\kappa \cdot 0}r(0) = \int_0^t e^{\kappa s} \kappa \theta ds + \int_0^t e^{\kappa s} \sigma dW(s) \Leftrightarrow \\
&\Leftrightarrow e^{\kappa t}r(t) = r(0) + \int_0^t e^{\kappa s} \kappa \theta ds + \int_0^t e^{\kappa s} \sigma dW(s) \Leftrightarrow \\
&\Leftrightarrow r(t) = e^{-\kappa t}r(0) + e^{-\kappa t} \int_0^t e^{\kappa s} \kappa \theta ds + e^{-\kappa t} \int_0^t e^{\kappa s} \sigma dW(s) \Leftrightarrow \\
&\Leftrightarrow r(t) = e^{-\kappa t}r(0) + \int_0^t e^{\kappa(s-t)} \kappa \theta ds + \int_0^t e^{\kappa(s-t)} \sigma dW(s).
\end{aligned}$$

□

Observación. Las integrales de la forma $\int_0^t \Delta(s)dW(s)$ se conocen como integrales de Itô. Como se demuestra en [11], estas integrales se distribuyen como una Normal con media 0 y varianza $\int_0^t \Delta^2(s)ds$. En nuestro caso, la integral $\int_0^t e^{\kappa(s-t)}\sigma dW(s)$ que aparece en la expresión de $r(t)$ es una integral de Itô y, por ello, conocemos su distribución.

El hecho que acabamos de comentar en la observación permite derivar otra de las propiedades del modelo de Vasicek:

Proposición 2. En el modelo de Vasicek, $r(t)$ se distribuye como una Normal con las siguientes funciones de media y varianza para cada t :

- *Media:* $m(r(t)) \equiv m(t) = e^{-\kappa t}r(0) + \theta(1 - e^{-\kappa t})$.
- *Varianza:* $v(r(t)) \equiv v(t) = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})$.

Demostración. Empezaremos la demostración observando que, como

$$r(t) = e^{-\kappa t}r(0) + \int_0^t e^{\kappa(s-t)} \kappa \theta ds + \int_0^t e^{\kappa(s-t)} \sigma dW(s),$$

la única fuente de incertidumbre en $r(t)$ proviene de $\int_0^t e^{\kappa(s-t)}\sigma dW(s)$. Como esta es una integral de Itô, sabemos que:

$$\int_0^t e^{\kappa(s-t)}\sigma dW(s) \sim N\left(0, v(t) = \int_0^t e^{2\kappa(s-t)}\sigma^2 ds\right).$$

Si operamos en la integral de la varianza:

$$\begin{aligned}
\int_0^t e^{2\kappa(s-t)}\sigma^2 ds &= \sigma^2 e^{-2\kappa t} \int_0^t e^{2\kappa s} ds = \sigma^2 e^{-2\kappa t} \frac{1}{2\kappa} \int_0^t 2\kappa e^{2\kappa s} ds = \\
&= \sigma^2 e^{-2\kappa t} \frac{1}{2\kappa} (e^{2\kappa t} - 1) = \frac{\sigma^2}{2\kappa} (1 - e^{-2\kappa t}).
\end{aligned}$$

Por ello, se tiene que:

$$\int_0^t e^{\kappa(s-t)}\sigma dW(s) \sim N\left(0, v(t) = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})\right). \quad (2.6)$$

Además, dado t , la expresión $e^{-\kappa t}r(0) + \int_0^t \kappa \theta e^{\kappa(s-t)} ds$ es constante y toma el valor $e^{-\kappa t}r(0) + \theta e^{-\kappa t}(e^{\kappa t} - 1) = e^{-\kappa t}r(0) + \theta(1 - e^{-\kappa t})$. Por tanto, la distribución de $r(t)$ será como la que tenga la integral de Itô, es decir, una Normal. Veamos a continuación su media y varianza.

La media será la media de la integral de Itô (que es nula), a la que se le añadirá el término constante $e^{-\kappa t}r(0) + \theta(1 - e^{-\kappa t})$. Por ello, se tiene que $m(t) = e^{-\kappa t}r(0) + \theta(1 - e^{-\kappa t})$.

La varianza será, simplemente, la varianza de la integral de Itô. Esto es así porque la suma de términos constantes no altera el valor de la varianza. Entonces, se tiene que $v(t) = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})$. \square

Proposición 3. *En el modelo de Vasicek, la covarianza del tipo a corto plazo en los instantes t_1 y t_2 viene dada por la expresión:*

$$c(r(t_1), r(t_2)) \equiv c(t_1, t_2) = \frac{\sigma^2}{2\kappa} e^{-\kappa(t_1+t_2)} \cdot (e^{2\kappa \min(t_1, t_2)} - 1).$$

Para demostrar esta proposición nos ayudaremos de una propiedad de la integral de Itô, conocida como **isometría de Itô**. Esta propiedad afirma que, dada una integral de Itô $I(t) = \int_0^t \Delta(u) dW(u)$, se tiene que $\mathbb{E}[I^2(t)] = \mathbb{E}\left[\int_0^t \Delta^2(u) du\right]$.

Veamos la demostración de la proposición anterior:

Demostración. La covarianza entre el tipo a corto plazo en los instantes t_1 y t_2 , $c(t_1, t_2)$, viene dada por:

$$c(t_1, t_2) = \mathbb{E}[r(t_1)r(t_2)] - \mathbb{E}[r(t_1)]\mathbb{E}[r(t_2)] = \mathbb{E}[r(t_1)r(t_2)] - m(t_1) \cdot m(t_2).$$

Podemos obtener una expresión de $r(t)$ en función de su media para facilitar la notación en los cálculos siguientes:

$$\begin{aligned} r(t) &= e^{-\kappa t}r(0) + \int_0^t e^{\kappa(s-t)} \kappa \theta ds + \int_0^t e^{\kappa(s-t)} \sigma dW(s) \Leftrightarrow \\ &\Leftrightarrow r(t) = e^{-\kappa t}r(0) + \theta(1 - e^{-\kappa t}) + \int_0^t e^{\kappa(s-t)} \sigma dW(s) \Leftrightarrow \\ &\Leftrightarrow r(t) = m(t) + \int_0^t e^{\kappa(s-t)} \sigma dW(s). \end{aligned}$$

Desarrollemos ahora el término $\mathbb{E}[r(t_1)r(t_2)]$ que aparece en la covarianza:

$$\begin{aligned} \mathbb{E}[r(t_1)r(t_2)] &= \mathbb{E}\left[\left(m(t_1) + \int_0^{t_1} e^{\kappa(s-t_1)} \sigma dW(s)\right) \cdot \left(m(t_2) + \int_0^{t_2} e^{\kappa(s-t_2)} \sigma dW(s)\right)\right] \\ &= \mathbb{E}[m(t_1) \cdot m(t_2)] + \\ &\quad + m(t_1) \int_0^{t_2} e^{\kappa(s-t_2)} \sigma dW(s) + m(t_2) \int_0^{t_1} e^{\kappa(s-t_1)} \sigma dW(s) + \\ &\quad + \left(\int_0^{t_1} e^{\kappa(s-t_1)} \sigma dW(s)\right) \cdot \left(\int_0^{t_2} e^{\kappa(s-t_2)} \sigma dW(s)\right). \end{aligned}$$

Como la esperanza es un operador lineal y las integrales $\int_0^{t_1} e^{\kappa(s-t_1)} \sigma dW(s)$ y $\int_0^{t_2} e^{\kappa(s-t_2)} \sigma dW(s)$ son de Itô, cuya esperanza es nula, se tiene que:

$$\mathbb{E}[r(t_1)r(t_2)] = m(t_1) \cdot m(t_2) + \mathbb{E} \left[\left(\int_0^{t_1} e^{\kappa(s-t_1)} \sigma dW(s) \right) \cdot \left(\int_0^{t_2} e^{\kappa(s-t_2)} \sigma dW(s) \right) \right].$$

Pasemos ahora a calcular la esperanza del producto de integrales:

$$\begin{aligned} & \mathbb{E} \left[\left(\int_0^{t_1} e^{\kappa(s-t_1)} \sigma dW(s) \right) \cdot \left(\int_0^{t_2} e^{\kappa(s-t_2)} \sigma dW(s) \right) \right] = \\ & = \mathbb{E} \left[\sigma^2 e^{-\kappa(t_1+t_2)} \left(\int_0^{t_1} e^{\kappa s} dW(s) \right) \cdot \left(\int_0^{t_2} e^{\kappa s} dW(s) \right) \right] = \\ & = \sigma^2 e^{-\kappa(t_1+t_2)} \cdot \mathbb{E} \left[\left(\int_0^{t_1} e^{\kappa s} dW(s) \right) \cdot \left(\int_0^{t_2} e^{\kappa s} dW(s) \right) \right]. \end{aligned}$$

Para poder aplicar la isometría de Itô en este momento, necesitamos que las dos integrales anteriores tengan los mismos límites de integración. Supongamos, sin pérdida de generalidad, que $t_1 \leq t_2$ (el caso en que $t_2 \leq t_1$ es análogo). Entonces, dividamos la integral entre 0 y t_2 en dos integrales (una entre 0 y t_1 y otra entre t_1 y t_2):

$$\begin{aligned} & \mathbb{E} \left[\left(\int_0^{t_1} e^{\kappa s} dW(s) \right) \cdot \left(\int_0^{t_2} e^{\kappa s} dW(s) \right) \right] = \\ & = \mathbb{E} \left[\int_0^{t_1} e^{\kappa s} dW(s) \cdot \left(\int_0^{t_1} e^{\kappa s} dW(s) + \int_{t_1}^{t_2} e^{\kappa s} dW(s) \right) \right] = \\ & = \mathbb{E} \left[\left(\int_0^{t_1} e^{\kappa s} dW(s) \right)^2 + \int_0^{t_1} e^{\kappa s} dW(s) \cdot \int_{t_1}^{t_2} e^{\kappa s} dW(s) \right] = \\ & = \mathbb{E} \left[\left(\int_0^{t_1} e^{\kappa s} dW(s) \right)^2 \right] + \mathbb{E} \left[\int_0^{t_1} e^{\kappa s} dW(s) \cdot \int_{t_1}^{t_2} e^{\kappa s} dW(s) \right]. \end{aligned}$$

Aplicando en el primer sumando la isometría de Itô tenemos que:

$$\mathbb{E} \left[\left(\int_0^{t_1} e^{\kappa s} dW(s) \right)^2 \right] = \mathbb{E} \left[\int_0^{t_1} e^{2\kappa s} ds \right] = \int_0^{t_1} e^{2\kappa s} ds = \frac{1}{2\kappa} (e^{2\kappa t_1} - 1).$$

Para el segundo sumando, como los incrementos de los movimientos brownianos (representados por $dW(s)$ en una y otra integral) son independientes, las integrales también lo serán. Debido a esto, la esperanza del producto de las integrales es igual al producto de las esperanzas de las integrales:

$$\mathbb{E} \left[\int_0^{t_1} e^{\kappa s} dW(s) \cdot \int_{t_1}^{t_2} e^{\kappa s} dW(s) \right] = \mathbb{E} \left[\int_0^{t_1} e^{\kappa s} dW(s) \right] \cdot \mathbb{E} \left[\int_{t_1}^{t_2} e^{\kappa s} dW(s) \right] = 0,$$

pues $\mathbb{E} \left[\int_0^{t_1} e^{\kappa s} dW(s) \right] = 0$ por ser una integral de Itô.

Volviendo al cálculo de la covarianza y recopilando todos los resultados parciales que hemos obtenido, llegamos a lo siguiente:

$$\begin{aligned} c(t_1, t_2) &= \mathbb{E}[r(t_1)r(t_2)] - m(t_1) \cdot m(t_2) = \\ &= m(t_1) \cdot m(t_2) + \sigma^2 e^{-\kappa(t_1+t_2)} \cdot \mathbb{E} \left[\left(\int_0^{t_1} e^{\kappa s} dW(s) \right) \cdot \left(\int_0^{t_2} e^{\kappa s} dW(s) \right) \right] - m(t_1) \cdot m(t_2) = \\ &= \sigma^2 e^{-\kappa(t_1+t_2)} \cdot \left(\frac{1}{2\kappa} (e^{2\kappa t_1} - 1) + 0 \right) = \frac{\sigma^2}{2\kappa} e^{-\kappa(t_1+t_2)} \cdot (e^{2\kappa t_1} - 1). \end{aligned}$$

Es decir, si $t_1 \leq t_2$, entonces:

$$c(t_1, t_2) = \frac{\sigma^2}{2\kappa} e^{-\kappa(t_1+t_2)} \cdot (e^{2\kappa t_1} - 1).$$

En el caso en que $t_2 \leq t_1$, obtendríamos:

$$c(t_1, t_2) = \frac{\sigma^2}{2\kappa} e^{-\kappa(t_1+t_2)} \cdot (e^{2\kappa t_2} - 1).$$

De manera que una expresión generalizada para la covarianza del tipo a corto plazo en los instantes t_1 y t_2 es:

$$c(t_1, t_2) = \frac{\sigma^2}{2\kappa} e^{-\kappa(t_1+t_2)} \cdot (e^{2\kappa \min(t_1, t_2)} - 1).$$

□

Por tanto, se concluye que el tipo a corto plazo en el modelo de Vasicek se distribuye como una distribución Normal multivariante caracterizada en cada instante temporal, respectivamente, por el siguiente vector de medias y matriz de varianzas y covarianzas:

$$m(t) = e^{-\kappa t} r(0) + \theta(1 - e^{-\kappa t}) \quad \text{y} \quad c(t_1, t_2) = \frac{\sigma^2}{2\kappa} e^{-\kappa(t_1+t_2)} \cdot (e^{2\kappa \min(t_1, t_2)} - 1).$$

2.2. Una expresión para los precios de los bonos, $P(t, T)$

El modelo de Vasicek tiene otra implicación importante y es que nos permite calcular una expresión para los precios (a tiempo t) de los bonos cupón cero que vencen en T , $P(t, T)$. Para ello, usaremos la siguiente proposición:

Proposición 4. *En un mercado sin arbitraje en el que \mathbb{Q} es una probabilidad neutral al riesgo, se tiene que $P(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} \right]$.*

Demostración. Como el mercado no tiene arbitraje, por el Primer Teorema Fundamental de Valoración Financiera³ aplicado a los precios de los bonos, $P(t, T)$, y tomando como numerario la cuenta bancaria, B_t , existe una medida de probabilidad neutral al riesgo, \mathbb{Q} ,

³Se puede consultar en el anexo A.

tal que el proceso $P(t, T)/B_t$ es martingala bajo dicha medida (dicho de otra manera, los precios de los bonos descontados con la cuenta bancaria son martingalas bajo \mathbb{Q}). Es decir, se tiene que:

$$\frac{P(t, T)}{B_t} = \mathbb{E}_t^{\mathbb{Q}} \left[\frac{P(T, T)}{B_T} \right].$$

Como $B_t = e^{\int_0^t r(s) ds}$, entonces lo anterior es equivalente a:

$$e^{-\int_0^t r(s) ds} \cdot P(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_0^T r(s) ds} \cdot P(T, T) \right].$$

Al ser $P(T, T) = 1$, tenemos:

$$e^{-\int_0^t r(s) ds} \cdot P(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds - \int_0^t r(s) ds} \right].$$

El hecho de calcular $\mathbb{E}_t^{\mathbb{Q}}[\cdot]$ quiere decir que la información hasta tiempo t es conocida. Por ello, el término $e^{-\int_0^t r(s) ds}$ puede salir de la esperanza:

$$e^{-\int_0^t r(s) ds} \cdot P(t, T) = e^{-\int_0^t r(s) ds} \cdot \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} \right].$$

Simplificando en los dos lados de la ecuación obtenemos la relación que queríamos demostrar:

$$P(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} \right].$$

□

Para obtener una expresión cerrada para $P(t, T)$ vamos a comenzar, entonces, calculando $\int_t^T r(s) ds$ partiendo de la ecuación (2.5):

$$r(t) = e^{-\kappa t} r(0) + \int_0^t e^{-\kappa(t-s)} \kappa \theta ds + \int_0^t e^{-\kappa(t-s)} \sigma dW(s).$$

Integrando entre t y T (y teniendo en cuenta que la variable s que aparece a continuación no es la misma que la de la ecuación (2.5), la cual pasa a ser u), tenemos:

$$\begin{aligned} \int_t^T r(s) ds &= \int_t^T e^{-\kappa s} r(0) ds + \int_t^T \int_0^s e^{-\kappa(s-u)} \kappa \theta du ds + \int_t^T \int_0^s e^{-\kappa(s-u)} \sigma dW(u) ds \Leftrightarrow \\ \Leftrightarrow \int_t^T r(s) ds &= e^{-\kappa t} r(0) \int_t^T e^{-\kappa(s-t)} ds + \int_t^T \int_0^s e^{-\kappa(s-u)} \kappa \theta du ds + \int_t^T \int_0^s e^{-\kappa(s-u)} \sigma dW(u) ds. \end{aligned}$$

Cambiando el orden de integración mediante la expresión⁴ $\int_t^T \int_0^s \cdot du ds = \int_0^t \int_t^T \cdot ds du + \int_t^T \int_u^T \cdot ds du$, llegamos a:

$$\begin{aligned} \int_t^T r(s) ds &= e^{-\kappa t} r(0) \int_t^T e^{-\kappa(s-t)} ds + \left[\int_0^t \int_t^T e^{-\kappa(s-u)} \kappa \theta ds du + \int_t^T \int_u^T e^{-\kappa(s-u)} \kappa \theta ds du \right] + \\ &+ \left[\int_0^t \int_t^T e^{-\kappa(s-u)} \sigma ds dW(u) + \int_t^T \int_u^T e^{-\kappa(s-u)} \sigma ds dW(u) \right]. \end{aligned}$$

⁴Esta expresión se comprueba fácilmente integrando todos los miembros.

Realizando algunos cambios de variable convenientes y operando, obtenemos:

$$\begin{aligned} \int_t^T r(s) ds &= \int_0^{T-t} e^{-\kappa s} ds \left[e^{-\kappa t} r(0) + \int_0^t e^{-\kappa(t-s)} \kappa \theta ds + \int_0^t e^{-\kappa(t-s)} \sigma dW(s) \right] + \\ &+ \int_t^T \left[\int_0^{T-u} e^{-\kappa s} ds \right] \kappa \theta du + \int_t^T \left[\int_0^{T-u} e^{-\kappa s} ds \right] \sigma dW(u). \end{aligned}$$

Si identificamos⁵ $E(t) \equiv \int_0^t e^{-\kappa s} ds$, entonces llegamos a la siguiente expresión:

$$\int_t^T r(s) ds = E(T-t)r(t) + \int_t^T E(T-u)\kappa\theta du + \int_t^T E(T-u)\sigma dW(u).$$

La integral $\int_t^T E(T-u)\sigma dW(u)$ (que es la única fuente de incertidumbre en la expresión anterior) se puede convertir en una integral de Itô haciendo un cambio de variable que transforme el dominio de integración en $[0, T-t]$. Entonces, procediendo como en la demostración de la proposición 2 de la sección anterior, concluimos que $\int_t^T r(s) ds$ se distribuye como una Normal con media $E(T-t)r(t) + \int_0^{T-t} E(u)\kappa\theta du$ y varianza $\int_0^{T-t} E^2(u)\sigma^2 du$.

Observación. Si Y es una variable aleatoria $N(\mu, \sigma)$, entonces su función generatriz de momentos es $M_Y(z) = \mathbb{E}[e^{zY}] = \exp(z\mu + \frac{1}{2}z^2\sigma^2)$.

Por la observación anterior, denotando $R \equiv \int_t^T r(s) ds$, tenemos:

$$M_R(z) = \mathbb{E}\left[e^{z \int_t^T r(s) ds}\right] = \exp\left(z \left[E(T-t)r(t) + \int_0^{T-t} E(u)\kappa\theta du \right] + \frac{1}{2}z^2 \int_0^{T-t} E^2(u)\sigma^2 du\right).$$

Recordemos que estamos buscando calcular $P(t, T) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r(s) ds}\right]$, por lo que haciendo $z = -1$ en la expresión anterior obtenemos el valor de $P(t, T)$:

$$\begin{aligned} P(t, T) &= \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r(s) ds}\right] = M_R(-1) = \\ &= \exp\left(- \left[E(T-t)r(t) + \int_0^{T-t} E(u)\kappa\theta du \right] + \frac{1}{2} \int_0^{T-t} E^2(u)\sigma^2 du\right). \end{aligned}$$

Como hemos denotado $E(t) = \int_0^t e^{-\kappa s} ds$, si ahora hacemos $D(t) = \int_0^t \left(E(u)\kappa\theta - E^2(u)\frac{\sigma^2}{2} \right) du$, obtenemos la expresión para el precio a tiempo t del bono que vence en T :

$$\boxed{P(t, T) = e^{-D(T-t) - E(T-t)r(t)}} \quad (2.7)$$

Esto quiere decir que en el modelo de Vasicek los precios de los bonos son Lognormales.

⁵No confundir con la notación para la esperanza, \mathbb{E} .

Sería interesante conseguir expresiones que no involucren integrales para $E(t)$ y $D(t)$. En el caso de $E(t)$, basta con integrar para obtener:

$$E(t) = \frac{1 - e^{-\kappa t}}{\kappa}.$$

Para $D(t)$, escribiendo la función en forma diferencial llegamos a $D'(t) = \kappa\theta E(t) - \frac{\sigma^2}{2}E^2(t)$. Haciendo algunos cálculos obtenemos que:

$$D(t) = \left(\theta - \frac{\sigma^2}{2\kappa^2}\right)(t - E(t)) + \frac{\sigma^2 E^2(t)}{4\kappa}.$$

Obviamente, las expresiones que hemos obtenido para D y E verifican que $D(0) = 0$ y $E(0) = 0$.

La expresión (2.7) para el precio de los bonos es de gran utilidad. Esto es debido a que las cantidades $P(t, T)$ son los factores de descuento que nos permiten obtener el valor, en el momento t , de una cantidad de dinero en el futuro (a tiempo T). Además, podemos expresar la mayoría de los pagos de los instrumentos financieros de renta fija en función de los precios de los bonos. En el siguiente capítulo nos encontraremos con estas situaciones en la práctica.

Capítulo 3

La metodología de Entropía Relativa Mínima (ERM) y Weighted Monte Carlo (WMC)

Presentamos en este capítulo el concepto de entropía y cómo da lugar a la metodología de Entropía Relativa Mínima, que constituye la parte central de este Trabajo. Veremos también cómo podemos aplicarla, mediante una técnica que se conoce como Monte Carlo Ponderado, para obtener una medida martingala equivalente a otra medida dada pero que, además, incorpore la presencia del smile en los precios de mercado.

3.1. El concepto de entropía relativa

La noción de entropía deriva del concepto de información. Formalmente, la información es una función, $I(p)$, asociada a la ocurrencia de un determinado suceso con una probabilidad p de tener lugar. Intuitivamente, esta función trata de medir la *cantidad de información* que proporciona la ocurrencia de dicho suceso bajo la medida de probabilidad que estamos usando. Por ejemplo, supongamos que tenemos acciones de una determinada compañía y se estima que mañana subirán con una probabilidad muy alta. Si mañana estas acciones suben no nos veremos sorprendidos, pues es precisamente lo que esperábamos el día anterior. Por este motivo, la ocurrencia de este suceso no nos proporcionará mucha información. Por el contrario, si las acciones bajan, este hecho generará mucha información, ya que era algo con muy poca probabilidad de ocurrir. Por tanto, la función de información proporcionada por la ocurrencia de un suceso que tenía probabilidad p de ocurrir, $I(p)$, será decreciente y no negativa.

A lo largo de los años se han propuesto diferentes funciones para cuantificar esta información. Por ejemplo, una elección frecuente para I es $I(p) = -\log(p)$, con $p \in (0, 1]$ por ser una probabilidad.

La función $-\log(p)$ es de la forma:

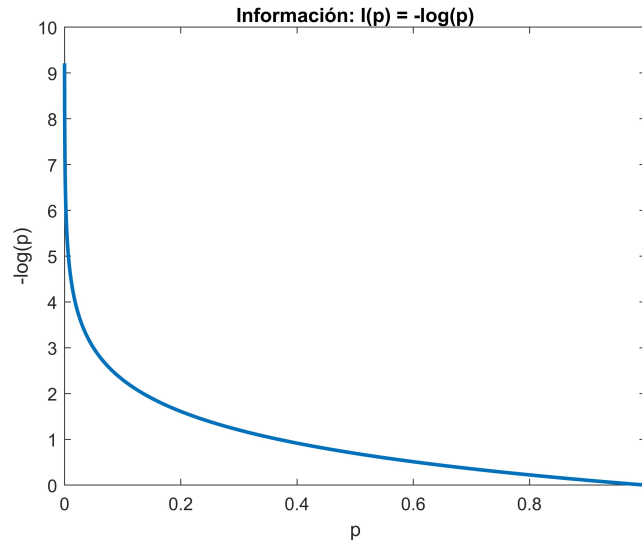


Figura 3.1: Función $I(p) = -\log(p)$

Como se puede ver en la figura anterior, $I(p)$ así definida es decreciente y no negativa. Además, refleja perfectamente el hecho de que, a menor probabilidad de ocurrencia de un suceso, mayor será la información que proporcionará este en caso de tener lugar. Si p está cerca de 0, entonces la información tiende a $+\infty$, mientras que si p se acerca a 1, la información tiende a ser nula.

Veamos ahora el concepto de entropía debido a C. E. Shannon:

Definición 2 (Entropía de una variable aleatoria discreta). *Consideremos una variable aleatoria discreta, X . Sea p_i la probabilidad de ocurrencia del suceso i -ésimo de X , con $i = 1, \dots, n$. Entonces, la entropía de X es una función, H , de la forma:*

$$H(X) = - \sum_{i=1}^n p_i \cdot \log(p_i). \quad (3.1)$$

Como $H(X) = - \sum_{i=1}^n p_i \cdot \log(p_i) = \sum_{i=1}^n p_i \cdot I(p_i)$, observamos que la entropía de una variable aleatoria X es el valor esperado de la información que proporciona la ocurrencia de los sucesos de X . Por este motivo, la entropía es una medida de la incertidumbre de la variable aleatoria X . Para ilustrar mejor este concepto, veamos un sencillo ejemplo:

Ejemplo 2. *Supongamos que la variable aleatoria X es degenerada, es decir, que el suceso x_1 aglutina toda la probabilidad de ocurrencia ($p_1 = 1$). Por este motivo, no hay incertidumbre a la hora de obtener una realización de X : sabemos que el resultado será el suceso*

x_1 . Dicho de otra manera, la información que nos proporciona la ocurrencia de este suceso es nula. Si calculamos la entropía de X , tenemos $H(X) = -1 \cdot \log(1) = 0$.

Pensemos ahora en otra variable aleatoria, Y , que asigne igual probabilidad a todos los sucesos posibles. Es decir, $P(Y = y_i) = \frac{1}{n}$, $\forall i = 1 \dots n$. Ahora, al hacer una realización de Y la incertidumbre será máxima, ya que todos los sucesos tienen la misma probabilidad de ocurrir. Al calcular la entropía de Y , obtenemos $H(Y) = -\sum_{i=1}^n \frac{1}{n} \cdot \log\left(\frac{1}{n}\right) = -\frac{1}{n} \sum_{i=1}^n -\log(n) = \log(n)$.

Podemos concluir, entonces, que a mayor (menor) incertidumbre, mayor (menor) información y, por tanto, mayor (menor) entropía. Además, del ejemplo podemos derivar cuáles son los valores entre los que se mueve la entropía de una variable aleatoria:

- El mayor valor que puede tomar la entropía de una variable aleatoria X es $\log(n)$. Se alcanza cuando los n sucesos de X son equiprobables.
- La entropía de la variable X es no negativa, es decir, $H(X) \geq 0$. Por tanto, el menor valor que puede tener es 0, el cual tiene lugar en el caso en que toda la probabilidad se concentra en un único suceso.

Supongamos ahora otra situación en la que tenemos dos variables aleatorias discretas, X e Y , y queremos obtener una forma de medir la “distancia” o diferencia que hay entre las distribuciones de estas (P^1 y P^2 , respectivamente). En el sentido de la inferencia estadística, podemos pensar que P^1 es una función de distribución real, mientras que P^2 es una estimación o aproximación de P^1 . En este caso, estamos interesados en cuantificar la *ineficiencia* al trabajar con la distribución aproximada, P^2 , en vez de con la exacta, P^1 . De esta manera aparece el concepto de entropía relativa de una función de distribución con respecto de otra (según el enfoque proporcionado por Kullback y Leibler, recogido en [5]):

Definición 3 (Entropía relativa de una función de distribución con respecto de otra). Sean dos variables aleatorias discretas, X e Y , con funciones de distribución¹ P y Q , respectivamente. Denotemos por p_i y q_i , con $i = 1, \dots, n$, a las probabilidades de ocurrencia de los sucesos i -ésimos de las variables X e Y , respectivamente. Entonces, la entropía relativa de P con respecto de Q se representa por $H(P/Q)$ y toma el valor:

$$H(P/Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i}.$$

La entropía relativa también se conoce como *distancia de Kullback-Leibler*. Conviene aclarar que, aunque intuitivamente podemos verla como una distancia entre las dos distribuciones, no lo es en sentido estricto ya que no es simétrica ni se cumple la desigualdad triangular en general.

¹Cambiamos la notación para adaptarla a la que usaremos en el resto del capítulo.

3.2. Entropía Relativa Mínima y simulación Monte Carlo: Monte Carlo Ponderado

En esta sección vamos a explicar la metodología de Entropía Relativa Mínima aplicada a la simulación Monte Carlo. Esta técnica se conoce como Monte Carlo Ponderado o *Weighted Monte Carlo* (WMC) y es la herramienta que emplearemos para tratar el problema del smile de volatilidad. Comenzamos motivando su uso y aplicación en este ámbito.

3.2.1. Idea general de la estrategia a seguir

El problema que estudiamos en este Trabajo es la necesidad de incorporar la influencia del smile de volatilidad en los modelos financieros. Dado el papel que desempeña la simulación Monte Carlo en este ámbito, la estrategia que seguiremos estará encaminada a considerar el smile de manera implícita en las simulaciones. Como es natural, nos plantearemos también cuál es la magnitud del riesgo de modelo existente en caso de no realizar este ajuste. Para ello, nos ayudaremos de un instrumento financiero, el swap, que depende de los tipos de interés.

Al realizar simulación Monte Carlo para valorar un determinado instrumento financiero y obtener métricas de exposición, se está suponiendo de manera implícita que todas las simulaciones tienen la misma probabilidad de ocurrir. Es decir, aparece una distribución (que denotaremos *a priori*) la cual asigna el mismo peso a todos los posibles escenarios en los que valoramos el swap y que intervendrán en el cálculo de las exposiciones (capítulo 4). Esta probabilidad *a priori*, al ser la probabilidad natural con la que ponderar las simulaciones de Monte Carlo, no incorpora la influencia del smile de volatilidad. Como ya comentamos en la introducción, la idea para dar solución a este problema consistía en tener más en cuenta la posible ocurrencia de eventos extremos. En este contexto, esto se traduce en modificar ligeramente la distribución con la que trabajamos, es decir, la distribución *a priori*. De esta manera, aparece la distribución que conoceremos como *a posteriori*, que recoge implícitamente la influencia del smile y permite dar distinto peso a cada escenario simulado.

De lo que acabamos de comentar se deduce la existencia de dos medidas de probabilidad, una *a priori*, Q , y otra *a posteriori*, P , que son equivalentes, aunque la probabilidad *a posteriori* es una extensión de la *a priori* ya que lleva implícito el efecto del smile de volatilidad:

- La medida Q es la distribución que asigna a todas las simulaciones el mismo peso, por lo que no incorpora el smile de volatilidad. Es una medida neutral al riesgo y constituye la distribución de partida en el método de Monte Carlo Ponderado.
- La medida P es la distribución que tiene en cuenta el smile de volatilidad. La obtendremos mediante la aplicación del método de Monte Carlo Ponderado y, al igual que Q , es una medida neutral riesgo.

Esta metodología, entonces, parte de un modelo calibrado a mercado que valora bien la curva de cupones cero (el modelo en que todas las simulaciones se ponderan con la probabilidad *a priori*) y permite obtener un valorador más completo. En nuestro caso, este será una probabilidad *a posteriori* que valora correctamente tanto los cupones cero como

algunas opciones *out the money*, cuyo precio se ve afectado por el smile de volatilidad. Las dos medidas, al valorar bien la curva cupón cero, serán **medidas martingalas equivalentes** y darán el mismo precio al instrumento que estamos valorando (en este caso, un swap). Sin embargo, el uso de la probabilidad a posteriori se verá reflejado en los valores de la exposición (como veremos más adelante), lo cual nos permitirá estudiar el riesgo de modelo (es decir, el riesgo asociado a no considerar la existencia del smile de volatilidad).

Las dos distribuciones de probabilidad serán muy parecidas y se diferencian en que P , como ya hemos comentado, otorga más peso a los eventos extremos (es más leptocúrtica que Q). Por tanto, la solución que proponemos para el problema planteado pasa por hacer un cambio de distribución a otra más completa, pero de manera que la *diferencia* entre las dos distribuciones sea pequeña. Esto motiva la consideración de la entropía relativa de P con respecto de Q tal y como la planteamos en la definición 3:

$$H(P/Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i}. \quad (3.2)$$

Como Q es la distribución que asigna la misma probabilidad a todos los eventos, $1/n$, tanto la información como la entropía de su variable aleatoria serán máximas. Por otro lado, al ser P más estrecha, la información y la entropía que nos proporcionará serán menores que las de Q .

Concretemos las condiciones que se tienen que verificar al hacer el cambio de Q a P :

- La distribución P que obtengamos mediante Monte Carlo Ponderado ha de ser muy parecida a Q . Por ello, la entropía relativa entre la distribución a posteriori y la distribución a priori ha de ser lo más pequeña posible. Esto se traduce en un problema de minimización de la entropía relativa de P con respecto de Q .
- Dispondremos de algunos *instrumentos financieros de referencia* cuyo precio de mercado (con smile incluido) conocemos a día de hoy. Por ello, pediremos que se valoren correctamente al usar la nueva distribución, P , junto con su pay-off descontado.
- La distribución a posteriori, P , ha de ser una medida martingala. Como se pone de manifiesto en [6], esta condición no siempre se verifica de manera natural al usar el método de Monte Carlo Ponderado, por lo que tenemos que exigirla adicionalmente.

Denotemos por $P(t, T)$, como es habitual, al precio a tiempo t del bono cupón cero que vence en T . Para clarificar la notación en este punto, vamos a denotar momentáneamente por \mathbf{P} a la medida a posteriori. Sabemos que esta medida \mathbf{P} es martingala si y solo si valora bien la curva cupón cero inicial, es decir, cuando se tiene que

$$\frac{P(0, T)}{B(0)} = \mathbb{E}_0^{\mathbf{P}} \left[\frac{P(T, T)}{B(T)} \right]$$

para todos los vencimientos T .

Como $B(0) = 1$ y $P(T, T) = 1$, lo anterior se traduce en:

$$P(0, T) = \mathbb{E}_0^{\mathbf{P}} \left[\frac{1}{B(T)} \right] = \mathbb{E}_0^{\mathbf{P}} \left[e^{-\int_0^T r(s) ds} \right]. \quad (3.3)$$

Por tanto, la condición exacta para que \mathbf{P} sea martingala se reduce a que se verifique (3.3) para un continuo de vencimientos posibles, T . En nuestro caso, vamos a introducir esta condición en el problema ayudándonos de los *Floating Rate Note* (FRN), pues son instrumentos muy líquidos en el mercado. A continuación veremos cómo hacerlo, aunque adelantamos que la condición de martingala se satisface cuando exigimos que se valoren correctamente los FRN en la medida \mathbf{P} , para todos los posibles vencimientos T .

El pago de un FRN en el instante t_{n+1} , asumiendo que el nominal es 1, viene dado por $L(t_n, t_n, t_{n+1}) \cdot \Delta t$. Al descontarlo, nos queda $P(0, t_{n+1}) \cdot L(t_n, t_n, t_{n+1}) \cdot \Delta t$. Como se tiene que

$$L(t_n, t_n, t_{n+1}) = \frac{1}{t_{n+1} - t_n} \left[\frac{P(t_n, t_n)}{P(t_n, t_{n+1})} - 1 \right] = \frac{1}{\Delta t} \left[\frac{1}{P(t_n, t_{n+1})} - 1 \right],$$

entonces el pago en t_{n+1} descontado del FRN es:

$$P(0, t_{n+1}) \left[\frac{1}{P(t_n, t_{n+1})} - 1 \right] = \left[\frac{P(0, t_{n+1})}{P(t_n, t_{n+1})} - P(0, t_{n+1}) \right] = [P(0, t_n) - P(0, t_{n+1})].$$

Entonces, si \mathbf{P} valora bien la curva cupón cero inicial para todos los vencimientos, también valorará bien la cantidad $[P(0, t_n) - P(0, t_{n+1})]$. Dicho de otra manera, cuando \mathbf{P} valora correctamente los cupones cero iniciales, también valora correctamente los FRN. Por ello, para garantizar que la medida \mathbf{P} obtenida mediante Monte Carlo Ponderado sea martingala, tendremos que introducir nuevas restricciones que aseguren que todos los FRN se valoren correctamente. Es decir, añadiremos los *Floating Rate Note* a los instrumentos de referencia.

Conviene notar que esta condición se cumple de manera exacta cuando valoramos los FRN para todos los vencimientos posibles. Como no podemos introducir un continuo de restricciones en el problema, tan solo exigiremos esta condición para los mismos vencimientos que tienen el resto de instrumentos de referencia. De esta manera, tendremos una condición aproximada que, como se verá posteriormente en la práctica, resulta suficiente para obtener buenos resultados.

De manera gráfica, el método que vamos a seguir se podría esquematizar en los cuatro bloques que se ven en la siguiente figura. Una de las ventajas de esta metodología es la **modularidad**, es decir, que tanto el bloque de la distribución a priori como el de las restricciones se pueden modificar sin afectar al resto. Este hecho es útil si, por ejemplo, encontramos una distribución a priori más adecuada (que se parezca más a la distribución real, P) o si deseamos añadir o cambiar restricciones.

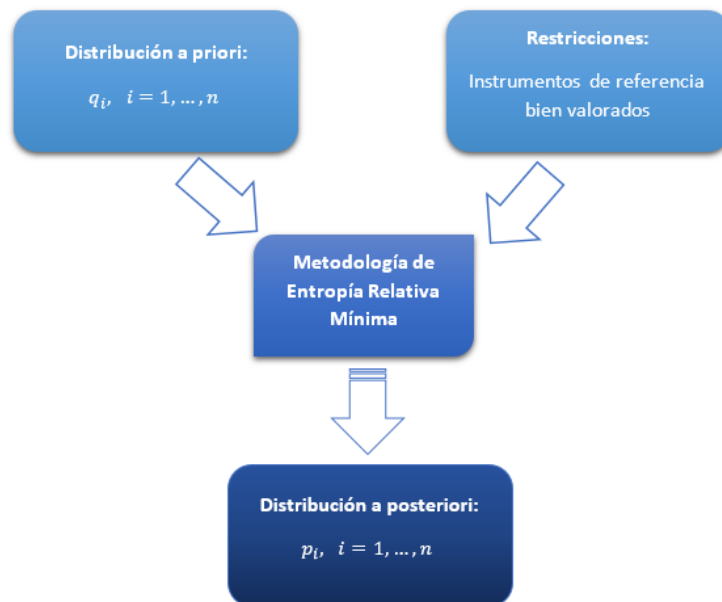


Figura 3.2: Esquema de la estrategia a seguir. Modificado de Arrieta D., [2].

Como vamos a ver a continuación, plantearemos un problema de minimización de la entropía relativa de la distribución a posteriori con respecto a la distribución a priori sujeto a restricciones. Al ser un problema de optimización, pueden darse los casos en que no se encuentre solución o en que esta no sea única. Si no obtenemos ninguna solución, entonces será debido a que no existe una probabilidad neutral al riesgo que cumpla las restricciones, es decir, que valore correctamente los instrumentos de referencia que hemos establecido. En base al Primer Teorema Fundamental de Valoración Financiera, esto pondría de manifiesto la existencia de arbitraje en el mercado. En cambio, si la solución no es única, por el Segundo Teorema Fundamental de Valoración Financiera² deduciríamos que el mercado no es completo, es decir, que existen contratos financieros que no se pueden replicar mediante otros activos del mercado.

3.2.2. Desarrollo teórico

Pasemos ahora a desarrollar de manera teórica el problema de optimización a resolver. Dado que la aplicación práctica de la metodología de Entropía Relativa Mínima que vamos a emplear en este Trabajo se conoce como Monte Carlo Ponderado, nos limitaremos a describir esta técnica. Realmente, esta es un caso particular de la metodología de ERM cuando:

- Las distribuciones de probabilidad involucradas son discretas (cada posible estado del conjunto de estados Ω^3 se identificará con una simulación, la cual se ponderará con la probabilidad correspondiente).
- La distribución a priori es equiprobable para cada estado (asigna la misma probabilidad de ocurrencia a todas las simulaciones).

² Ambos teoremas se pueden consultar en el anexo A.

³ En este caso, Ω será el conjunto de posibles estados del mercado.

En este marco, la función de distribución a posteriori que vamos a obtener nos servirá para dar más peso a determinadas simulaciones (intuitivamente, a las que se correspondan con estados del mercado más extremos y desfavorables, pues son a las que queremos dar más probabilidad de ocurrencia). Por este motivo, al usarla estaremos teniendo en cuenta de manera implícita la influencia del smile de volatilidad.

Comencemos introduciendo la notación que vamos a seguir:

Notación.

- n representa el número de simulaciones que vamos a realizar, es decir, el número de posibles estados del mercado que consideramos.
- m es el número de instrumentos de referencia de los que conocemos su precio y que formarán parte de las restricciones del problema de optimización.
- El índice i , con $i = 1, \dots, n$, hará referencia a la i -ésima simulación.
- El índice j , con $j = 1, \dots, m$, denotará al instrumento de referencia j -ésimo.
- $\mathbf{q} = (q_1, \dots, q_n)$ y $\mathbf{p} = (p_1, \dots, p_n)$ son los vectores de probabilidades⁴ a priori y a posteriori, respectivamente, de los posibles estados del mercado. Por ello, al resolver el problema de optimización habremos determinado el vector \mathbf{p} .
- Los elementos $g_{i,j}$ denotan el valor actual (descontado con el valor de la cuenta bancaria) del instrumento j en el estado del mercado i . Además, \mathbf{g}_j denotará el vector columna de precios descontados del instrumento j en todas las simulaciones. Llamaremos \mathbf{G} a la matriz cuyas columnas son estos vectores: $\mathbf{G} = \{\mathbf{g}_j, j = 1, \dots, m\} = \{g_{i,j}, i = 1, \dots, n, j = 1, \dots, m\}$.
- El valor C_j denota el precio actual en el mercado del instrumento j (es un valor con smile incluido que se obtiene del mercado).

Como acabamos de comentar, los precios C_j , para $j = 1, \dots, m$, de los instrumentos de referencia ya llevan incluido el smile. Recordemos que las restricciones de nuestro problema exigirán que la probabilidad a posteriori, $\mathbf{p} = (p_1, \dots, p_n)$, que minimice la entropía relativa proporcione el valor correcto a los precios simulados de estos instrumentos (que sean iguales a los de mercado, con una determinada tolerancia). Si existe dicha probabilidad a posteriori, como será una probabilidad neutral al riesgo, por el Primer Teorema Fundamental de Valoración Financiera tendremos garantizado que el mercado está libre de oportunidades de arbitraje. En ese caso, teniendo en cuenta de nuevo este teorema y tomando como nominal la cuenta bancaria, las restricciones serán de la forma:

$$\sum_{i=1}^n p_i \cdot g_{i,j} = C_j, \quad j = 1, \dots, m. \quad (3.4)$$

Como $\sum_{i=1}^n p_i \cdot g_{i,j} = \mathbb{E}^{\mathbf{P}}[\mathbf{g}_j]$, se dice que este método incorpora restricciones en los momentos de la distribución a posteriori.

⁴Recalamos que, en el caso general, \mathbf{q} se corresponde con la función de densidad de la distribución de probabilidad a priori y \mathbf{p} con la de la distribución a posteriori.

Los instrumentos de referencia que tomaremos serán opciones *out the money* sobre los tipos de interés (*caplets* y *floorlets*, como veremos en el capítulo dedicado a la resolución práctica del problema). Los precios de mercado de estos instrumentos serán, por tanto, más elevados que los que se obtienen con la fórmula de Black-Scholes para sus respectivos strikes. Debemos de ser cuidadosos a la hora de determinar qué instrumentos seleccionamos para aplicar la técnica de Monte Carlo Ponderado, es decir, cuando establecemos qué strikes vamos a considerar para los *caplets* y *floorlets out the money*. Si consideramos strikes asociados a opciones extremadamente *out the money*, muchos elementos de la matriz \mathbf{G} serán nulos (o estarán cerca de serlo) y esto puede traer problemas a la hora de aplicar la técnica que describiremos a continuación. Por ello, debemos realizar una **calibración** previa para determinar los strikes de las opciones que vamos a incluir, de manera que estas estén *out the money* en determinados escenarios, pero que en otros sí tengan valor y proporcionen información útil al algoritmo de resolución del problema. También nos puede ayudar en esta selección el hecho de que, cuando un *floorlet* tiene muchos escenarios OTM, entonces el *caplet* con el mismo strike tendrá todos esos escenarios ITM (por la paridad entre los *caplets* y *floorlets*). Por ello, si el instrumento seleccionado es un *caplet* o un *floorlet* muy OTM, bastaría con cambiarlo por el contrario, manteniendo el mismo strike.

Resumiendo todo lo comentado anteriormente: tenemos que plantear un problema de minimización de la entropía relativa de la distribución a posteriori (\mathbf{p}) con respecto a la distribución a priori ($q_i = 1/n$, $i = 1, \dots, n$) (ecuación (3.2)), sujeto a las restricciones dadas por (3.4) y, también, imponiendo que \mathbf{p} sea una medida de probabilidad (es decir, que sus elementos sumen 1). De esa manera, obtendremos los valores para p_i , $i = 1, \dots, n$, los cuales constituyen una probabilidad neutral al riesgo que incorpora de manera implícita la influencia del smile. Lo que estaremos haciendo es un cambio en la medida martingala, pasando de \mathbf{q} a \mathbf{p} .

La formulación del problema de optimización es, por tanto, la siguiente:

$$\begin{array}{l} \underset{\mathbf{p}}{\text{mín}} \quad H(\mathbf{p}/\mathbf{q}) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i} \\ \text{s. a:} \quad \sum_{i=1}^n p_i \cdot g_{i,j} = C_j, \quad j = 1, \dots, m \\ \quad \quad \quad \sum_{i=1}^n p_i = 1 \end{array}$$

Para resolver este problema de optimización con restricciones vamos a usar el método de los multiplicadores de Lagrange. Como se ve en [4], formulando el dual del problema con restricciones pasaremos a tener un problema de optimización minimax en el que cada una

de las restricciones originales se incluirá en la función objetivo junto con un peso, λ_j :

$$\min_{\lambda_0, \mathbf{\Lambda}} \left[\max_{\mathbf{p}} \left\{ - \sum_{i=1}^n p_i \log \frac{p_i}{q_i} + \sum_{j=1}^m \lambda_j \left[\sum_{i=1}^n p_i \cdot g_{i,j} - C_j \right] + \lambda_0 \cdot \left(\sum_{i=1}^n p_i - 1 \right) \right\} \right]. \quad (3.5)$$

Comenzamos resolviendo el problema de maximización. Para ello, fijamos la cantidad λ_0 y el vector $\mathbf{\Lambda} = (\lambda_1, \dots, \lambda_m)$ y buscamos las probabilidades \mathbf{p} que maximizan el Lagrangiano para esos λ_0 y $\mathbf{\Lambda}$. Derivando con respecto de cada p_i e igualando a 0:

$$\begin{aligned} \frac{\partial}{\partial p_i} \left[- \sum_{i=1}^n \left(\overbrace{p_i \log \frac{p_i}{q_i}}^{p_i(\log p_i - \log q_i)} \right) + \sum_{j=1}^m \lambda_j \left[\sum_{i=1}^n p_i \cdot g_{i,j} - C_j \right] + \lambda_0 \cdot \left(\sum_{i=1}^n p_i - 1 \right) \right] &= 0 \Leftrightarrow \\ \Leftrightarrow - \left[\log p_i - \log q_i + p_i \cdot \left(\frac{1}{p_i} \right) \right] + \sum_{j=1}^m \lambda_j \cdot g_{i,j} + \lambda_0 &= 0 \Leftrightarrow \\ \Leftrightarrow - \log p_i + \log q_i - 1 + \sum_{j=1}^m \lambda_j \cdot g_{i,j} + \lambda_0 &= 0 \Leftrightarrow \\ \Leftrightarrow p_i = \exp \left(\log q_i + \sum_{j=1}^m \lambda_j \cdot g_{i,j} + \lambda_0 - 1 \right) &\Leftrightarrow \\ \Leftrightarrow p_i = q_i \cdot \exp \left(\sum_{j=1}^m \lambda_j \cdot g_{i,j} \right) \cdot e^{(\lambda_0 - 1)}. & \quad (3.6) \end{aligned}$$

Entonces, fijados λ_0 y el vector $\mathbf{\Lambda}$, las probabilidades óptimas son las dadas por (3.6) para $i = 1, \dots, n$. Veamos que esta solución es, efectivamente, un máximo. Si volvemos a derivar con respecto de cada p_i , obtenemos que:

$$\frac{\partial}{\partial p_i} \left[- \log p_i + \log q_i - 1 + \sum_{j=1}^m \lambda_j \cdot g_{i,j} + \lambda_0 \right] = -\frac{1}{p_i}.$$

Como p_i es positivo, se tiene que la segunda derivada es negativa y, por tanto, p_i definido como en (3.6) es un máximo.

Ahora sustituiremos las probabilidades que acabamos de obtener en (3.6) en la expresión del problema que estamos resolviendo, (3.5). De esta manera, llegamos al problema de minimización en λ_0 y $\mathbf{\Lambda}$:

$$\min_{\lambda_0, \mathbf{\Lambda}} \left[- \sum_{i=1}^n p_i \log \frac{p_i}{q_i} + \sum_{j=1}^m \lambda_j \left[\sum_{i=1}^n p_i \cdot g_{i,j} - C_j \right] + \lambda_0 \cdot \left(\sum_{i=1}^n p_i - 1 \right) \right] =$$

$$\begin{aligned}
 &= \min_{\lambda_0, \Lambda} \left[- \sum_{i=1}^n p_i \cdot \overbrace{\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j} + \lambda_0 - 1 \right)}^{\log(p_i/q_i)} + \sum_{j=1}^m \lambda_j \left[\sum_{i=1}^n p_i \cdot g_{i,j} - C_j \right] + \lambda_0 \cdot \left(\sum_{i=1}^n p_i - 1 \right) \right] = \\
 &= \min_{\lambda_0, \Lambda} \left[- \sum_{i=1}^n \sum_{j=1}^m p_i \lambda_j g_{i,j} - \sum_{i=1}^n p_i (\lambda_0 - 1) + \sum_{j=1}^m \sum_{i=1}^n \lambda_j p_i g_{i,j} - \sum_{j=1}^m \lambda_j C_j + \lambda_0 \left(\sum_{i=1}^n p_i - 1 \right) \right] = \\
 &= \min_{\lambda_0, \Lambda} \left[- \sum_{i=1}^n p_i \cdot (\lambda_0 - 1) - \sum_{j=1}^m \lambda_j \cdot C_j + \lambda_0 \cdot \left(\sum_{i=1}^n p_i - 1 \right) \right] = \\
 &= \min_{\lambda_0, \Lambda} \left[\sum_{i=1}^n p_i - \sum_{j=1}^m \lambda_j \cdot C_j - \lambda_0 \right] = \\
 &= \min_{\lambda_0, \Lambda} \left[\sum_{i=1}^n \overbrace{\left(q_i \cdot \exp \left(\sum_{j=1}^m \lambda_j \cdot g_{i,j} \right) \cdot e^{(\lambda_0-1)} \right)}^{p_i} - \sum_{j=1}^m \lambda_j \cdot C_j - \lambda_0 \right]. \quad (3.7)
 \end{aligned}$$

Minimicemos esta función derivando con respecto de cada λ_j , para $j = 0, 1, \dots, m$, e igualando a cero. Empecemos con el caso en que $j \neq 0$:

$$\begin{aligned}
 &\frac{\partial}{\partial \lambda_j} \left[\sum_{i=1}^n \left(q_i \cdot \exp \left(\sum_{j=1}^m \lambda_j \cdot g_{i,j} \right) \cdot e^{(\lambda_0-1)} \right) - \sum_{j=1}^m \lambda_j \cdot C_j - \lambda_0 \right] = 0 \Leftrightarrow \\
 &\Leftrightarrow \sum_{i=1}^n \left(q_i \cdot g_{i,j} \cdot \exp \left(\sum_{j=1}^m \lambda_j \cdot g_{i,j} \right) \cdot e^{(\lambda_0-1)} \right) - C_j = 0 \Leftrightarrow \sum_{i=1}^n p_i \cdot g_{i,j} - C_j = 0.
 \end{aligned}$$

Veamos el caso con λ_0 :

$$\begin{aligned}
 &\frac{\partial}{\partial \lambda_0} \left[\sum_{i=1}^n \left(q_i \cdot \exp \left(\sum_{j=1}^m \lambda_j \cdot g_{i,j} \right) \cdot e^{(\lambda_0-1)} \right) - \sum_{j=1}^m \lambda_j \cdot C_j - \lambda_0 \right] = 0 \Leftrightarrow \\
 &\Leftrightarrow \sum_{i=1}^n \left(q_i \cdot \exp \left(\sum_{j=1}^m \lambda_j \cdot g_{i,j} \right) \cdot e^{(\lambda_0-1)} \right) - 1 = 0 \Leftrightarrow \sum_{i=1}^n p_i - 1 = 0.
 \end{aligned}$$

De los desarrollos anteriores no podemos obtener una expresión para los λ_j , con $j = 0, 1, \dots, m$, puesto que en ambos hemos obtenido de nuevo las restricciones de nuestro problema. Por ello, necesitaremos encontrar otro método para llegar a la solución del problema. Antes de ello, hagamos un inciso y veamos que la última condición implica que:

$$\begin{aligned}
e^{(\lambda_0-1)} \cdot \sum_{i=1}^n q_i \cdot \exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right) &= 1 \Leftrightarrow \\
\Leftrightarrow e^{(\lambda_0-1)} &= \frac{1}{\sum_{i=1}^n q_i \cdot \exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)}. \tag{3.8}
\end{aligned}$$

Sustituyendo este valor de $e^{(\lambda_0-1)}$ en la expresión de p_i (ecuación (3.6)), obtenemos:

$$p_i = \frac{q_i \cdot \exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)}{\sum_{i=1}^n q_i \cdot \exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)}, \quad i = 1, \dots, n. \tag{3.9}$$

Al escribir de esta manera las cantidades p_i podremos obtener expresiones más compactas para nuestros desarrollos.

En este momento, antes de seguir estudiando el problema de minimización, podemos sustituir las cantidades q_i por las probabilidades que hemos elegido. Como ya hemos comentado, en nuestro caso la distribución a priori será la que asigna a todas las simulaciones el mismo peso⁵, es decir, $q_i = 1/n$, $i = 1, \dots, n$. Esto nos permite trabajar con una expresión de p_i simplificada:

$$p_i = \frac{\exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)}{\sum_{i=1}^n \exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)}, \quad i = 1, \dots, n. \tag{3.10}$$

Denotemos por $Z(\mathbf{\Lambda})$ a la cantidad $\sum_{i=1}^n \exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)$. De esta manera observamos que $Z(\mathbf{\Lambda})$ es un *factor normalizador*, pues obliga a que la suma de todos los p_i sea 1 (como ha de ocurrir, pues \mathbf{p} tiene que ser una función de densidad). Entonces, podemos

reescribir la expresión (3.10) como $p_i = \frac{\exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)}{Z(\mathbf{\Lambda})}$, para $i = 1, \dots, n$.

Volvamos de nuevo al problema de minimización de la expresión (3.7). Haciendo $q_i = 1/n$, $\forall i$, en la ecuación (3.8), podemos obtener el valor de λ_0 :

$$e^{(\lambda_0-1)} = \frac{n}{Z(\mathbf{\Lambda})} \Leftrightarrow \lambda_0 - 1 = \log n - \log Z(\mathbf{\Lambda}) \Leftrightarrow \lambda_0 = \log n - \log Z(\mathbf{\Lambda}) + 1.$$

Como vemos, este valor quedará determinado cuando conozcamos los elementos del vector $\mathbf{\Lambda}$. Teniendo esto en cuenta y, de nuevo, particularizando para el caso en que $q_i = 1/n$, $\forall i$, el problema de minimización dado por (3.7) se transforma en:

$$\min_{\mathbf{\Lambda}} \left[\sum_{i=1}^n \frac{\exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)}{\sum_{i=1}^n \exp\left(\sum_{j=1}^m \lambda_j \cdot g_{i,j}\right)} - \sum_{j=1}^m \lambda_j \cdot C_j \overbrace{-\log n + \log Z(\mathbf{\Lambda}) - 1}^{-\lambda_0} \right] =$$

⁵Por ello, la entropía relativa para esta distribución a priori en particular es $H(\mathbf{p}/\mathbf{q}) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i} = \sum_{i=1}^n p_i (\log p_i + \log n) = \log n + \sum_{i=1}^n p_i \log p_i$, suponiendo que $\sum_{i=1}^n p_i = 1$ al ser \mathbf{p} una probabilidad.

$$\begin{aligned}
 &= \underset{\mathbf{\Lambda}}{\text{mín}} \left[\frac{1}{Z(\mathbf{\Lambda})} \sum_{i=1}^n \exp \left(\overbrace{\sum_{j=1}^m \lambda_j \cdot g_{i,j}}^{Z(\mathbf{\Lambda})} \right) - \sum_{j=1}^m \lambda_j \cdot C_j - \log n + \log Z(\mathbf{\Lambda}) - 1 \right] = \\
 &= \underset{\mathbf{\Lambda}}{\text{mín}} \left[- \sum_{j=1}^m \lambda_j \cdot C_j - \log n + \log Z(\mathbf{\Lambda}) \right]. \tag{3.11}
 \end{aligned}$$

Como $\log n$ es una constante que no afectará a la hora de obtener los valores $(\lambda_1, \dots, \lambda_m)$ que minimizan la expresión (3.11), el problema de minimización en $\mathbf{\Lambda}$ es, realmente:

$$\underset{\mathbf{\Lambda}}{\text{mín}} \left[\log Z(\mathbf{\Lambda}) - \sum_{j=1}^m \lambda_j \cdot C_j \right]. \tag{3.12}$$

Denotemos por $W(\mathbf{\Lambda}) = \log Z(\mathbf{\Lambda}) - \sum_{j=1}^m \lambda_j C_j$ a la función a minimizar (es simplemente (3.7) particularizada en el caso en que $q_i = 1/n, \forall i$). Como ya vimos, la condición de gradiente nulo ($\nabla W(\mathbf{\Lambda}) = \mathbf{0}$) se obtiene por cálculo directo y es equivalente a que se cumplan las restricciones en los momentos de la distribución a posteriori:

$$\nabla W(\mathbf{\Lambda}) = \begin{pmatrix} \frac{\partial W(\mathbf{\Lambda})}{\partial \lambda_1} \\ \vdots \\ \frac{\partial W(\mathbf{\Lambda})}{\partial \lambda_m} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n p_i g_{i,1} - C_1 \\ \vdots \\ \sum_{i=1}^n p_i g_{i,m} - C_m \end{pmatrix} = \begin{pmatrix} \mathbb{E}^{\mathcal{P}}[\mathbf{g}_1] - C_1 \\ \vdots \\ \mathbb{E}^{\mathcal{P}}[\mathbf{g}_m] - C_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Lo anterior solo nos confirma que los precios obtenidos mediante este método para los instrumentos de referencia son, efectivamente, consistentes con los precios de mercado de los mismos. Sin embargo, dado que no podemos deducir una expresión para cada λ_j del desarrollo anterior, en la literatura se pueden encontrar algoritmos que realizan esta tarea y nos permiten resolver el problema dado por (3.12). En nuestro caso, nos ayudaremos del algoritmo propuesto en [6] para encontrar el vector $\mathbf{\Lambda}$ que minimiza $W(\mathbf{\Lambda})$.

Antes de describirlo, comprobemos que la solución al problema dado por (3.12) es, efectivamente, un mínimo. Para ello, denotemos por \mathbf{H} a la matriz hessiana de $W(\mathbf{\Lambda})$, es decir, a la matriz que recoge las segundas derivadas de $W(\mathbf{\Lambda})$ con respecto a los distintos λ_j . Veamos cómo es el elemento en la fila u y en la columna v de \mathbf{H} :

$$\mathbf{H}(u, v) = \frac{\partial^2 W(\mathbf{\Lambda})}{\partial \lambda_u \partial \lambda_v} = \frac{\partial}{\partial \lambda_v} \left(\overbrace{\frac{\partial W(\mathbf{\Lambda})}{\partial \lambda_u}}^{\mathbb{E}^{\mathcal{P}}[\mathbf{g}_u] - C_u} \right) = \frac{\partial \mathbb{E}^{\mathcal{P}}[\mathbf{g}_u]}{\partial \lambda_v} = \frac{\partial}{\partial \lambda_v} \left(\sum_{i=1}^n p_i \cdot g_{i,u} \right) =$$

Introduciendo el valor de p_i , para $i = 1, \dots, n$:

$$\begin{aligned} &= \frac{\partial}{\partial \lambda_v} \left(\frac{1}{Z(\boldsymbol{\Lambda})} \sum_{i=1}^n g_{i,u} \exp \left(\sum_{j=1}^m \lambda_j g_{i,j} \right) \right) = \\ &= \frac{-1}{Z(\boldsymbol{\Lambda})^2} \frac{\partial Z(\boldsymbol{\Lambda})}{\partial \lambda_v} \cdot \sum_{i=1}^n g_{i,u} \exp \left(\sum_{j=1}^m \lambda_j g_{i,j} \right) + \underbrace{\frac{1}{Z(\boldsymbol{\Lambda})}} \cdot \sum_{i=1}^n \underbrace{g_{i,u} g_{i,v} \exp \left(\sum_{j=1}^m \lambda_j g_{i,j} \right)} = \end{aligned}$$

De nuevo identificando p_i para $i = 1, \dots, n$:

$$\begin{aligned} &= \frac{-1}{Z(\boldsymbol{\Lambda})} \frac{\partial Z(\boldsymbol{\Lambda})}{\partial \lambda_v} \cdot \sum_{i=1}^n g_{i,u} \underbrace{\frac{1}{Z(\boldsymbol{\Lambda})} \exp \left(\sum_{j=1}^m \lambda_j g_{i,j} \right)} + \sum_{i=1}^n g_{i,u} g_{i,v} p_i = \\ &= \frac{-1}{Z(\boldsymbol{\Lambda})} \frac{\partial Z(\boldsymbol{\Lambda})}{\partial \lambda_v} \cdot \sum_{i=1}^n g_{i,u} p_i + \sum_{i=1}^n g_{i,u} g_{i,v} p_i = \end{aligned}$$

$$\begin{aligned} \text{Como } \frac{\partial Z(\boldsymbol{\Lambda})}{\partial \lambda_v} &= \sum_{i=1}^n g_{i,v} \exp \left(\sum_{j=1}^m \lambda_j g_{i,j} \right): \\ &= - \underbrace{\frac{1}{Z(\boldsymbol{\Lambda})}} \sum_{i=1}^n \left(\underbrace{g_{i,v} \exp \left(\sum_{j=1}^m \lambda_j g_{i,j} \right)} \right) \cdot \sum_{i=1}^n g_{i,u} p_i + \sum_{i=1}^n g_{i,u} g_{i,v} p_i = \end{aligned}$$

Volvemos a agrupar el término p_i :

$$= - \sum_{i=1}^n g_{i,v} p_i \cdot \sum_{i=1}^n g_{i,u} p_i + \sum_{i=1}^n g_{i,u} g_{i,v} p_i = \mathbb{E}^{\mathbf{P}}[\mathbf{g}_u \cdot \mathbf{g}_v] - \mathbb{E}^{\mathbf{P}}[\mathbf{g}_u] \cdot \mathbb{E}^{\mathbf{P}}[\mathbf{g}_v] = \text{Cov}^{\mathbf{P}}(\mathbf{g}_u, \mathbf{g}_v).$$

Es decir, $\mathbf{H}(u, v) = \text{Cov}^{\mathbf{P}}(\mathbf{g}_u, \mathbf{g}_v)$, $\forall (u, v)$. De esta forma, concluimos que la matriz hessiana de $W(\boldsymbol{\Lambda})$, \mathbf{H} , es la matriz de varianzas y covarianzas con respecto a la probabilidad a posteriori, \mathbf{p} , de los vectores de precios descontados de los instrumentos de referencia (\mathbf{g}_j). Por este motivo, la matriz \mathbf{H} es semidefinida positiva⁶ y, por tanto, $W(\boldsymbol{\Lambda})$ es convexa (equivalentemente: el punto crítico de $W(\boldsymbol{\Lambda})$ es un mínimo). En el caso de que los vectores de precios descontados, \mathbf{g}_j , sean linealmente independientes⁷, la matriz \mathbf{H} será definida

⁶Sus autovalores son no negativos.

⁷Lo cual es deseable ya que, como se explica en [3], las redundancias en los instrumentos de referencia pueden llevar a la inestabilidad del método.

positiva⁸, por lo que $W(\mathbf{\Lambda})$ será estrictamente convexa y el mínimo será único.

Describamos ahora el algoritmo que vamos a seguir para encontrar el vector $\mathbf{\Lambda}$ que minimiza $W(\mathbf{\Lambda})$. Se trata de un método iterativo que se basa en el desarrollo de Taylor de orden 2 de la función $W(\mathbf{\Lambda})$ alrededor de un punto, $\hat{\mathbf{\Lambda}}$:

$$W(\mathbf{\Lambda}) \approx W(\hat{\mathbf{\Lambda}}) + \nabla W(\hat{\mathbf{\Lambda}})(\mathbf{\Lambda} - \hat{\mathbf{\Lambda}}) + \frac{1}{2}(\mathbf{\Lambda} - \hat{\mathbf{\Lambda}})^T \mathbf{H}(\mathbf{\Lambda} - \hat{\mathbf{\Lambda}}). \quad (3.13)$$

La estrategia propuesta por este algoritmo consiste en resolver el problema de programación cuadrática sin restricciones que aparece al hacer el desarrollo de Taylor de orden 2 para el punto inicial, $\hat{\mathbf{\Lambda}}$. Este problema aproximado tiene solución analítica, la cual podremos usar para determinar otro nuevo punto inicial y repetir el algoritmo de manera iterativa. En cada iteración se comprobará el gradiente obtenido con la solución parcial y se parará cuando este sea nulo o esté cerca de serlo con una tolerancia dada.

La condición de optimalidad (gradiente nulo) para este problema se obtiene calculando el gradiente de la aproximación de $W(\mathbf{\Lambda})$ dada en (3.13):

$$\begin{aligned} \frac{\partial}{\partial \mathbf{\Lambda}} (W(\hat{\mathbf{\Lambda}}) + \nabla W(\hat{\mathbf{\Lambda}})(\mathbf{\Lambda} - \hat{\mathbf{\Lambda}}) + \frac{1}{2}(\mathbf{\Lambda} - \hat{\mathbf{\Lambda}})^T \mathbf{H}(\mathbf{\Lambda} - \hat{\mathbf{\Lambda}})) &= 0 \Leftrightarrow \\ \Leftrightarrow \nabla W(\hat{\mathbf{\Lambda}}) + \mathbf{H}(\mathbf{\Lambda} - \hat{\mathbf{\Lambda}}) &= \mathbf{0}. \end{aligned}$$

Podemos despejar $\mathbf{\Lambda}$ de la ecuación anterior para obtener la solución parcial del problema (es decir, la solución del problema aproximado de programación cuadrática). Como este valor no tiene por qué cumplir las condiciones de gradiente nulo del problema original (dado por (3.12)), para ser estrictos con la notación lo denotaremos por $\mathbf{\Lambda}_1$:

$$\mathbf{\Lambda}_1 = \hat{\mathbf{\Lambda}} - \mathbf{H}^{-1} \nabla W(\hat{\mathbf{\Lambda}}). \quad (3.14)$$

La estrategia consiste en, como ya anticipamos, realizar iteraciones de los pasos que acabamos de describir (fijando para la siguiente iteración $\hat{\mathbf{\Lambda}} = \mathbf{\Lambda}_1$) hasta que obtengamos un valor de $\mathbf{\Lambda}_1$ que verifique que $\nabla W(\mathbf{\Lambda}_1) \approx 0$, al menos con una determinada tolerancia. Sin embargo, si obtenemos la solución aproximada del problema como en (3.14), es probable que se produzcan grandes variaciones al pasar de una a otra, lo que no es deseable si ya estamos cerca del punto con gradiente nulo del problema original. La solución pasa por la introducción de un factor α que disminuya el tamaño del salto entre dos soluciones. Como se detalla en [6], un valor de inicio posible para este factor es $\alpha = 0.01$, el cual se duplicará en cada iteración hasta llegar a un máximo de $\alpha = 1$. Por ello, el paso de una solución parcial a otra se llevará a cabo mediante la ecuación:

$$\mathbf{\Lambda}_1 = \hat{\mathbf{\Lambda}} - \alpha \mathbf{H}^{-1} \nabla W(\hat{\mathbf{\Lambda}}). \quad (3.15)$$

Cuando $\nabla W(\mathbf{\Lambda}_1) \approx 0$ (con $\mathbf{\Lambda}_1$ dado como en (3.15)), el algoritmo parará y tomaremos $\mathbf{\Lambda}^* = \mathbf{\Lambda}_1$. De no ser así, la siguiente iteración se hará tomando como punto inicial $\hat{\mathbf{\Lambda}} = \mathbf{\Lambda}_1$.

⁸Se puede probar con facilidad que el rango de la matriz de varianzas y covarianzas (con respecto a la probabilidad a posteriori) de los vectores de la matriz \mathbf{G} es igual al rango de dicha matriz \mathbf{G} . Por ello, si los elementos de \mathbf{G} son linealmente independientes, la matriz \mathbf{G} tendrá rango máximo (es decir, m), al igual que la matriz \mathbf{H} . Es decir, \mathbf{H} no tendrá ningún autovalor nulo.

En el momento en que tengamos el vector $\mathbf{\Lambda}^*$ que minimiza (3.12), habremos resuelto el problema. Calcularemos el factor de normalización de las probabilidades a posteriori, $Z(\mathbf{\Lambda}^*) = \sum_{i=1}^n \exp\left(\sum_{j=1}^m \lambda_j^* g_{i,j}\right)$, y, a continuación, dichas probabilidades (ecuación (3.10)):

$$p_i = \frac{1}{Z(\mathbf{\Lambda}^*)} \exp\left(\sum_{j=1}^m \lambda_j^* g_{i,j}\right).$$

Conviene tener en cuenta que este algoritmo puede tener problemas de funcionamiento en algunos casos:

- Como comentamos al plantear el problema, si los instrumentos de referencia están muy *out the money*, la matriz \mathbf{G} tendrá muchos elementos nulos. Por tanto, la matriz hessiana, definida mediante $\mathbf{H}(u, v) = \text{Cov}^{\mathcal{P}}(\mathbf{g}_u, \mathbf{g}_v)$, $\forall (u, v)$, tendrá elementos muy cercanos a cero, lo que puede dar problemas a la hora de calcular su inversa.
- Es conveniente que los vectores de precios descontados, \mathbf{g}_j , sean linealmente independientes. De no ser así, la matriz \mathbf{H} no será definida positiva y, nuevamente, podemos tener problemas en el cálculo de la inversa⁹.
- Si el mercado no está libre de oportunidades de arbitraje, el algoritmo no converge. Este caso se da cuando existen opciones para cubrir una situación concreta en un determinado espacio temporal que tienen distintos strikes pero el mismo precio. Habrá un conflicto a la hora de verificar la condición óptima de gradiente nulo, $\mathbb{E}^{\mathcal{P}}[\mathbf{g}_j] = C_j$, ya que tanto \mathbf{p} como C_j serán iguales en esos casos, mientras que \mathbf{g}_j no.

En resumen, los pasos que hemos seguido para resolver el problema de la obtención de probabilidades que contengan de manera implícita la influencia del smile de volatilidad son:

1. Generamos n trayectorias del tipo a corto plazo con el modelo de Vasicek. Cada una de ellas tiene la misma probabilidad de ocurrir que las demás, por lo que se establece la distribución a priori como $q_i = \frac{1}{n}$, para $i = 1, \dots, n$.
2. Seleccionamos de manera justificada los instrumentos de referencia de los que conocemos el precio de mercado con smile incluido. Además, también estableceremos los *Floating Rate Note*, FRN, que nos servirán para garantizar que la probabilidad a posteriori que calculemos sea martingala.
3. Calculamos la matriz de flujos (descontados) de los instrumentos de referencia y los FRN, \mathbf{G} . Los flujos de estos instrumentos dependerán de las trayectorias del tipo a corto plazo generadas en el primer paso.
4. Pasamos a realizar la optimización del problema. Para ello, fijaremos el valor de los multiplicadores de Lagrange, $\mathbf{\Lambda}$, al vector nulo¹⁰ como punto de partida inicial y aplicaremos el algoritmo descrito.

⁹Esto es debido a que puede tener algún autovalor nulo y, como el determinante de una matriz es el producto de sus autovalores, este también será nulo.

¹⁰Es decir, comenzaremos resolviendo el problema sin restricciones. Esto es equivalente a que la probabilidad a posteriori sea igual que la a priori.

5. Cuando hayamos obtenido el vector de multiplicadores de Lagrange óptimos, $\mathbf{\Lambda}^*$, calcularemos las probabilidades a posteriori asociadas a él. Entonces, el problema estará resuelto y dispondremos de unas probabilidades a posteriori que recogen de manera implícita la influencia del smile de volatilidad. A partir de ahora, para valorar con smile usaremos estas probabilidades en vez de las a priori.

Capítulo 4

El riesgo de modelo cuantificado mediante métricas de exposición

Las métricas de exposición son medidas del valor que tiene para nosotros, a lo largo del tiempo, una determinada operación financiera (cartera, instrumento...) en la que participamos. Dado que no se puede calcular de manera exacta la exposición real que tenemos (puesto que no conocemos de manera determinista cómo va a evolucionar el mercado), estas métricas son estadísticas que cuantifican cuánto nos deberá la contrapartida en el caso de incurrir en impago (hecho conocido como *hacer default*). Por ello, permiten gestionar el riesgo de *default* de la contraparte: si esta quiebra, la cantidad de dinero que nosotros perderemos será el valor en ese instante de la operación financiera que manteníamos con ella.

Sea V el precio de nuestra cartera o *portfolio*. Denotemos $V^+ = \max(V, 0)$ y $V^- = \min(V, 0)$. Entonces, las métricas de exposición más utilizadas se definen como sigue:

- **Exposición positiva esperada** (*Expected Positive Exposure, EPE_t*): Representa cuánto nos pueden deber, en media, en el instante t . Formalmente, si $\Psi_t(V)$ es la distribución de V en el instante t , entonces la exposición positiva esperada en t es:

$$EPE_t = \int_{-\infty}^{\infty} V^+ \cdot \Psi_t(V) \cdot dV.$$

Para aplicarla en el caso en que la distribución de V en t sea discreta (que es el caso que se dará en nuestra práctica, como veremos), para cada instante t nos quedaremos con los valores de la cartera positivos, haciendo 0 los no positivos. Después, ponderaremos todos con las probabilidades correspondientes a la distribución de V en t .

- **Exposición negativa esperada** (*Expected Negative Exposure, ENE_t*): Análogamente a la EPE_t , refleja cuánto podemos deber nosotros, en media, a la contrapartida en el momento t :

$$ENE_t = \int_{-\infty}^{\infty} V^- \cdot \Psi_t(V) \cdot dV.$$

Particularizándola de nuevo en el caso discreto, seleccionaremos los valores de la cartera negativos y haremos 0 los no negativos. A continuación, ponderaremos con las probabilidades asociadas a la distribución de V en t .

- **Exposición potencial futura** (*Potential Future Exposure, PFE_t*): Para emplear esta métrica fijaremos un nivel de confianza, $1 - \alpha$. La PFE_t es tal que la probabilidad de que el valor de la cartera o del instrumento a tiempo t caiga por debajo de la misma es del $(1 - \alpha)$ %.

Por ello, la exposición potencial futura en t es el valor PFE_t que verifica:

$$1 - \alpha = \int_{-\infty}^{PFE_t^{1-\alpha}} \Psi_t(V) \cdot dV.$$

Normalmente se toman niveles de confianza desde el 90 % hasta el 99 %, dependiendo del apetito por el riesgo que tenga cada institución financiera. Para realizar la práctica hemos tomado un nivel intermedio: el 95 %.

En el caso en que la distribución de V en t sea discreta, se trata de calcular el percentil 95 según dicha distribución de probabilidad. Para llevar a cabo este cálculo, comenzaremos ordenando de manera ascendente los precios de la cartera en t , ordenando de la misma manera las probabilidades de ocurrencia de cada uno de esos precios (para no perder la correspondencia). Después calcularemos la suma acumulada de dichas probabilidades de ocurrencia, de manera que obtengamos la distribución empírica asociada a ellas. Entonces, identificaremos qué precio corresponde a la probabilidad acumulada que sea igual a 0.95. Como esto en general no ocurrirá, se puede redondear o, mejor, interpolar linealmente. Para ello, obtendremos los precios cuya probabilidad acumulada esté inmediatamente por encima y por debajo de 0.95 y realizaremos la interpolación lineal, obteniendo así el valor para PFE_t .

Aunque no es puramente una métrica de exposición, en la práctica también vamos a calcular el precio esperado de la cartera en el instante t , conocido como **precio Mark to Market** en t , MtM_t . Este precio se calcula ponderando con los pesos adecuados los distintos valores que hemos obtenido en cada simulación para la cartera en el instante t . Además, el MtM_t se puede calcular de otra manera, ya que resulta ser la suma de la exposición positiva esperada y de la exposición negativa esperada:

$$\begin{aligned} MtM_t &= \int_{-\infty}^{\infty} V \cdot \Psi_t(V) \cdot dV = \int_{-\infty}^{\infty} (V^+ + V^-) \cdot \Psi_t(V) \cdot dV = \\ &= \int_{-\infty}^{\infty} V^+ \cdot \Psi_t(V) \cdot dV + \int_{-\infty}^{\infty} V^- \cdot \Psi_t(V) \cdot dV = EPE_t + ENE_t. \end{aligned}$$

Como la EPE_t y la ENE_t representan el mismo concepto pero entendido desde diferentes perspectivas (la ENE_t propia es la EPE_t que calcula la contrapartida y viceversa), esta descomposición hace que el Mark to Market en t explicita las exposiciones de ambas partes.

Capítulo 5

Valoración de un swap de tipos de interés con *smile* y análisis del riesgo de modelo

Como refleja el título del Trabajo, el objetivo último de este es valorar un swap sobre los tipos de interés poniendo de manifiesto el riesgo de modelo existente. Existen numerosos factores que influyen en la falta de adecuación de un modelo para reflejar la realidad que representa. Aquí nos centramos en uno de ellos: la no consideración de la existencia del *smile* de volatilidad en el mercado.

En este capítulo, en primer lugar, se llevará a cabo la valoración del swap en base al tipo de interés a corto plazo simulado con el modelo de Vasicek. Para ello, usaremos la probabilidad neutral al riesgo obtenida mediante la aplicación de Monte Carlo Ponderado, que refleja la influencia del *smile* de volatilidad. Después estudiaremos el riesgo de modelo mediante algunas métricas de exposición.

5.1. Formulación del problema a resolver

La situación que se presenta consiste en valorar un swap de tipos de interés en el que pagamos el interés fijo (K) y recibimos el variable (*swap payer*) sobre un nominal, \mathcal{N} , de 1000 euros. El swap comienza dentro de 6 meses¹ y tiene una duración de 1.5 años (es decir, acaba en $T = 2$). Los pagos son semestrales y se realizan al final de cada periodo. El interés fijo que pagamos es $K = 0.07$ con composición semestral y el interés variable es el Libor que se fija para cada periodo al principio del mismo.

Queremos realizar la valoración a lo largo de una malla diaria de tiempos para obtener posteriormente los perfiles de algunas métricas de exposición al riesgo de crédito. Como veremos, esta parte consiste en ver cuál es el valor del swap para nosotros en cada instante (si la contrapartida hiciera default, esa sería la cantidad que podríamos perder). Para estudiar el riesgo de emplear un modelo inadecuado en la valoración del swap, realizaremos esta en

¹A los *swaps* que comienzan en una fecha futura se les conoce como *forward start swaps*.

dos casos diferentes: cuando simulamos mediante Monte Carlo (dando el mismo peso a todas las simulaciones) y cuando lo hacemos por Monte Carlo Ponderado, dando más importancia a las observaciones extremas. Analizaremos los resultados obtenidos en ambos casos para obtener algunas conclusiones.

5.2. Simulación del tipo de interés a corto plazo mediante el modelo de Vasicek

Como todos los instrumentos que intervienen en este problema y en su resolución dependen de los tipos de interés, necesitaremos hacer simulaciones de $r(t)$ que nos sirvan como punto de partida para evaluarlos. Para ello, emplearemos el modelo de Vasicek que explicamos en el capítulo 2.

El primer paso consiste en fijar la partición de tiempos en los que deseamos conocer el valor de la short rate. En nuestro caso, haremos 10000 simulaciones de la short rate entre 0 y 2 años con saltos temporales diarios², es decir, $\tau = \left(\frac{1}{30}\right) \left(\frac{1}{12}\right)$. También se tienen que calibrar los parámetros del modelo, tarea que no es trivial y para la que hay investigación al respecto. En nuestro caso, supondremos que la calibración ya se ha realizado y que estos parámetros toman los siguientes valores:

$$\theta = 0.08, \kappa = 0.86, \sigma = 0.01.$$

Además, establecemos $r(0) = 0.06$.

Para simular el tipo de interés a corto plazo tenemos dos opciones:

- La primera consiste en partir del valor de $r(0)$. Para cada simulación, calcularemos $dr(t) = \kappa(\theta - r(t))dt + \sigma dW(t)$ para los tiempos de nuestra partición con $t > 0$ e iremos obteniendo cada $r(t)$ en función del anterior, incrementando este con el valor $dr(t)$ calculado.
- Otra opción consiste en simular de la variable aleatoria Normal multivariante caracterizada por el vector de medias $m(t) = e^{-\kappa t}r(0) + \theta(1 - e^{-\kappa t})$ y la matriz de varianzas y covarianzas $c(t_1, t_2) = \frac{\sigma^2}{2\kappa}e^{-\kappa(t_1+t_2)} \cdot (e^{2\kappa \min(t_1, t_2)} - 1)$.

En el Trabajo nos hemos decantado por la segunda opción, pero la primera hubiera sido igual de factible.

²Aunque en situaciones reales se tienen en cuenta convenciones para medir los tiempos en el ámbito financiero, en este Trabajo no las usaremos por simplicidad.

En la siguiente imagen se muestran 20 de las 10000 trayectorias de la short rate que hemos simulado:

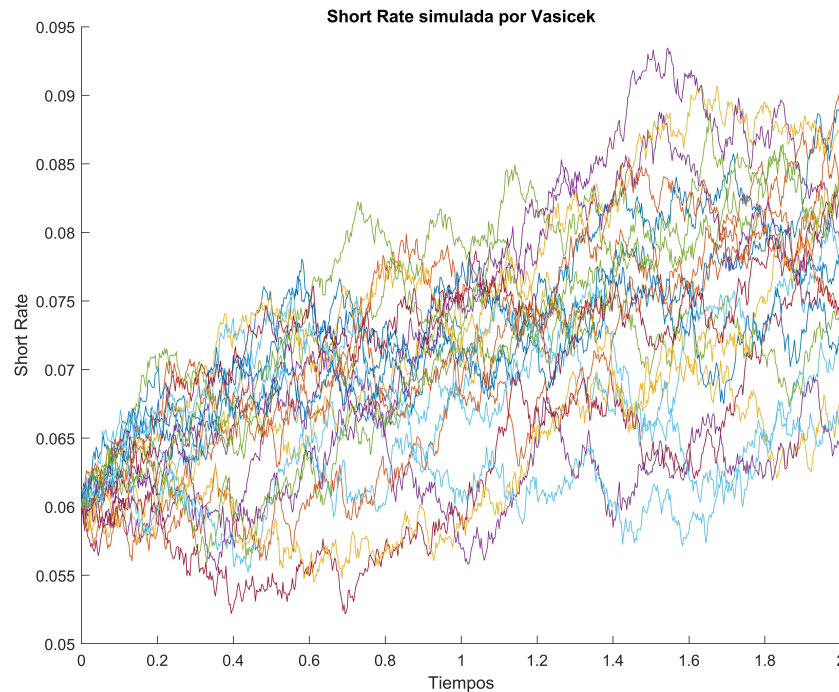


Figura 5.1: Algunas trayectorias simuladas del tipo a corto plazo.

Para poder trabajar de manera más eficiente, hemos guardado la matriz con las simulaciones del tipo a corto plazo en un archivo *.mat* que se importará en los siguientes scripts cuando sea necesario.

5.3. Obtención de una medida de probabilidad con *smile*

La problemática del smile de volatilidad, como comentamos anteriormente, se ha abordado desde diferentes perspectivas a lo largo de los años. La elegida por nosotros consiste en la modificación del peso que se asigna a cada simulación para valorar el swap, dando más importancia a los escenarios más extremos y desfavorables. De esa manera, obtendremos una probabilidad neutral al riesgo (*probabilidad a posteriori*) que integra de manera implícita la influencia del smile de volatilidad en los precios. Por ello, el uso de la probabilidad a posteriori ofrece ventajas a la hora de obtener resultados del modelo (como las métricas de exposición). En el capítulo 3 hemos desarrollado la estrategia teórica que permite la resolución de este problema y ahora pasamos a ponerla en práctica.

La primera decisión que debemos tomar para aplicar el método de Monte Carlo Ponderado hace referencia a los instrumentos de referencia que vamos a considerar. Esta elección debe estar justificada, pues dependiendo de los instrumentos que seleccionemos podemos

obtener solución o que el algoritmo no funcione. Ya anticipamos los casos en los que el algoritmo de optimización para obtener los multiplicadores, $\mathbf{\Lambda}$, no funciona correctamente. Recordémoslos:

- La matriz \mathbf{G} , al contener los valores descontados de los instrumentos de referencia seleccionados, puede tener muchos elementos nulos si estos están demasiado *out the money*. Por ello, debemos ser cuidadosos al seleccionar los strikes de las opciones que incluyamos. Se debe realizar una calibración previa en la que se observe, para cada posible strike de los instrumentos, qué porcentaje de simulaciones tienen valor estrictamente positivo. En caso de que este sea demasiado bajo, el algoritmo se *fijará* demasiado en estas (pocas) simulaciones, concentrando en esos escenarios toda la probabilidad a posteriori. Esto provocará un mal funcionamiento del mismo, pudiendo incluso no obtener ninguna solución.
- Si existen valores descontados de algunos instrumentos, \mathbf{g}_j , que sean linealmente dependientes, la matriz hessiana no será definida positiva. El algoritmo de optimización no funcionará al no poder calcular la inversa de \mathbf{H} . Por ello, debemos elegir instrumentos cuyos payoffs descontados sean linealmente independientes.
- Cuando el mercado no está libre de oportunidades de arbitraje. Este hecho, como es frecuente en finanzas, lo daremos por supuesto siempre.

Como en nuestro problema estamos considerando instrumentos dependientes de los tipos de interés, las opciones que se incluirán como instrumentos de referencia serán *caps* y *floors out the money*. Más concretamente, incluiremos los *caplets* y *floorlets*³ en los que se dividen estos *caps* y *floors* (que realmente son carteras de opciones). Por otro lado, como ya anticipamos anteriormente, también vamos a incluir los instrumentos conocidos como *Floating Rate Note* (FRN). Si algún *caplet* (*floorlet*) está demasiado *out the money*, se puede sustituir por un *floorlet* (*caplet*) del mismo strike, que estará *in the money* y no introducirá muchos elementos nulos en \mathbf{G} . En concreto, en este problema hemos incluido 3 *caplets* ITM y 3 OTM, 3 *floorlets* OTM y 3 FRN.

Las fechas de referencia para los instrumentos anteriores serán $0 \leq t_0 < t_1 < t_2 < \dots < t_N$. En $t_0 < t_1 < \dots < t_{N-1}$ se realizan fijaciones de los valores a pagar por el instrumento al final del periodo correspondiente (es decir, al inicio de cada periodo se fija el pago del instrumento al final del mismo). Por tanto, $t_1 < \dots < t_N$ son las fechas de pago.

Describamos estos instrumentos para obtener expresiones de su payoff:

- Los *caps* son instrumentos que aseguran una cota superior al interés. Como hemos anticipado, los *caps* están formados por los *caplets*, instrumentos más pequeños que representan una opción en cada intervalo temporal considerado. En cada fecha de fijación, t_n , se observa el tipo Libor que corresponde al intervalo que comienza en t_n , $L_n := L(t_n, t_n, t_{n+1})$. Por tanto, en t_{n+1} el *caplet* pagará:

$$\mathcal{N} \times \max(L_n - K, 0) \times (t_{n+1} - t_n).$$

³Son los instrumentos análogos a las funciones call y put, respectivamente, en el marco de los tipos de interés.

- Los *floors*, por el contrario, garantizan una cota inferior al interés. Se descomponen en los *floorlets*. De manera análoga al pago que se realizaba en los *caplets*, el *floorlet* pagará en t_{n+1} :

$$\mathcal{N} \times \max(K - L_n, 0) \times (t_{n+1} - t_n).$$

- Los *Floating Rate Note*, FRN, pagan el interés correspondiente a cada intervalo temporal cuando este acaba. Es decir, el pago (aleatorio hasta que se fija en t_n) de un FRN en t_{n+1} es:

$$\mathcal{N} \times L(t_n, t_n, t_{n+1}) \times (t_{n+1} - t_n).$$

Notemos que en estos instrumentos no se realizan pagos de nominal, solo del interés sobre dicho nominal.

En nuestro problema consideraremos que los instrumentos comienzan en tiempo 0, que realizan pagos semestrales y que tienen duración 2 años. Es decir, las fechas de fijación serán $t = 0, t = 0.5, t = 1$ y $t = 1.5$, mientras que las de pagos serán $t = 0.5, t = 1, t = 1.5$ y $t = 2$. Dado que cuando se ha fijado un pago este ya es conocido, los pagos de los instrumentos que fijan en $t = 0$ y pagan en $t = 0.5$ están determinados desde el comienzo. Por ello, no vamos a incluirlos en los instrumentos de referencia. Así, en el problema incluiremos solo los instrumentos que hacen fijaciones en $t_1 = 0.5, t_2 = 1$ y $t_3 = 1.5$.

A continuación, se debe realizar la calibración de los strikes para los *caplets* y los *floorlets*. Una vez realizada esta, hemos seleccionado los siguientes instrumentos de referencia:

Instrumento	Fecha de fijación	Fecha de pago	Strike	Porcentaje ITM
Caplet ITM 1	0.5	1	0.07	54.72 %
Caplet ITM 2	1	1.5	0.07	78.53 %
Caplet ITM 3	1.5	2	0.07	87.2 %
Caplet OTM 1	0.5	1	0.08	2.9 %
Caplet OTM 2	1	1.5	0.08	17.47 %
Caplet OTM 3	1.5	2	0.08	30.99 %
Floorlet OTM 1	0.5	1	0.0687	34.63 %
Floorlet OTM 2	1	1.5	0.0687	15.62 %
Floorlet OTM 3	1.5	2	0.0687	9.04 %
FRN 1	0.5	1	-	-
FRN 2	1	1.5	-	-
FRN 3	1.5	2	-	-

Cuadro 5.1: Instrumentos de referencia, strikes y porcentaje de simulaciones ITM.

Observemos que los *caplets* y *floorlets* de cada *cap* y *floor* tienen el mismo strike, que es el que vamos a calibrar para todos ellos. Sin embargo, están activos en distintos intervalos de tiempo y a cada uno le corresponde un Libor diferente. Por este motivo, las proporciones de los instrumentos con un mismo strike que están *in the money* varían dentro de cada *cap* o *floor*. Como vemos en la tabla anterior, dentro del *cap* y del *floor out the money*

encontramos proporciones que son suficientemente altas (a partir de un 20% ó 30% de las simulaciones, aproximadamente), mientras que otras son más bajas. Por tanto, tenemos que encontrar los strikes más equilibrados, en el sentido de que no dejen algunos *caplets* o *floorlets* muy *out the money*, mientras que otros del mismo *cap* o *floor* estén casi *in the money*. Haciendo varias pruebas en Matlab para controlar este hecho, finalmente nos hemos quedado con los que aparecen en la tabla superior. Además, hemos comprobado que con esta configuración los vectores \mathbf{g}_j , con $j = 1, \dots, 12$, son linealmente independientes, ya que la matriz \mathbf{G} tiene rango máximo.

Como estamos haciendo un ejemplo de aplicación de Monte Carlo Ponderado, no disponemos de datos de mercado reales. Por ello, vamos a calcular manualmente unos *precios de mercado* que incluyan el *smile* para estos instrumentos.

Empezamos obteniendo precios sin *smile* de volatilidad. Esto es equivalente a llevar a cabo 10000 simulaciones Monte Carlo usuales de los precios de los instrumentos y obtener el valor medio de todas ellas. Es decir, es equivalente a realizar un Monte Carlo ponderado en el que la distribución a posteriori asigna igual probabilidad a todas las simulaciones⁴.

Para incluir de manera manual el *smile*, tan solo tenemos que aumentar ligeramente los precios obtenidos de los *caplets* y *floorlets* que estaban *out the money*. Este aumento se puede realizar directamente, añadiendo arbitrariamente una pequeña cantidad a los precios sin *smile*. Sin embargo, esto es poco preciso y puede derivar en precios que no sean alcanzables ponderando las simulaciones con una medida neutral al riesgo. Una opción mejor consiste en calcular las volatilidades implícitas⁵ de los precios sin *smile* mediante la fórmula de Black para *caplets* y *floorlets* activos en el intervalo $[t_n, t_{n+1}]$:

$$V_0 = \mathcal{N}(t_{n+1} - t_n) \cdot P_0(t_{n+1}) \cdot \psi \cdot (L_0(t_n, t_{n+1}) \cdot \Phi(\psi d_1) - K \cdot \Phi(\psi d_2)),$$

donde:

$$\begin{aligned} v &:= \sigma_n \sqrt{t_n}, \\ d_1 &:= \frac{\log(L_0(t_n, t_{n+1})/K) + v^2/2}{v}, \\ d_2 &:= \frac{\log(L_0(t_n, t_{n+1})/K) - v^2/2}{v}, \end{aligned}$$

ψ toma valor 1 si se está valorando un *caplet* y -1 si se valora un *floorlet*, σ_n es la volatilidad de la variación continua del Libor subyacente en $[t_n, t_{n+1}]$ y Φ es la densidad acumulada de la distribución Normal estándar.

Cuando dispongamos de las volatilidades implícitas de esos precios, las añadiremos una pequeña constante (en la práctica hemos añadido un 0.55% de volatilidad a los *caplets* con strike 0.08 y un 0.5% a los *floorlets* con strike 0.0687). A continuación calcularemos de nuevo, mediante la fórmula de valoración para *caplets* y *floorlets* anterior, los precios de

⁴Este hecho significa que la distribución a posteriori es igual a la priori y ocurre cuando se plantea el problema de minimización sin restricciones. Dicho de otra manera, si no hay restricciones, entonces la entropía relativa mínima que se alcanza es 0, que corresponde al caso en que las dos distribuciones son iguales.

⁵La volatilidad implícita, como anticipamos al principio del Trabajo, es el valor de la volatilidad que hace que el precio obtenido con la fórmula de valoración sea igual al de mercado.

estos instrumentos con las nuevas volatilidades. Los nuevos precios serán más altos que los anteriores y, por ello, los consideraremos como precios de mercado con *smile* incluido para las opciones *out the money*. Recordemos que los precios de los *caplets in the money* y del FRN no se ven modificados, ya que no tienen *smile* de volatilidad.

En este punto, entonces, realizamos la optimización para minimizar la entropía relativa mínima, tal y como se describió en el capítulo 3. Comenzamos fijando $\mathbf{\Lambda} = (\lambda_1, \dots, \lambda_m) = (0, \dots, 0)$ y calculando las probabilidades a posteriori para ese vector, dadas por la expresión (3.10). Con ellas, obtenemos los precios mediante WMC ($\mathbb{E}^{\mathcal{P}}[g_j]$) y el gradiente de $W(\mathbf{\Lambda})$, $\nabla W(\mathbf{\Lambda}) = \mathbb{E}^{\mathcal{P}}[g_j] - C_j$. En ese momento comenzamos a realizar iteraciones del algoritmo descrito, actualizando el vector $\mathbf{\Lambda}$ mediante la ecuación (3.15), hasta conseguir que el gradiente esté suficientemente cerca de cero. La tolerancia que hemos establecido para aceptar un gradiente como nulo es 10^{-9} . Es decir, si el gradiente es menor que esta cantidad, consideraremos que hemos obtenido el vector $(\lambda_1, \dots, \lambda_m)$ que resuelve el problema de minimización (3.12). Entonces, aplicando (3.10), conseguimos las probabilidades a posteriori que incorporan de manera implícita la influencia del *smile* de volatilidad en los instrumentos OTM. Por tanto, ya podemos valorar y obtener resultados derivados (como las exposiciones) teniendo en cuenta esta característica.

5.3.1. Resultados obtenidos

Los precios sin *smile* y con *smile* (precios de mercado) que hemos obtenido para cada instrumento de referencia son:

Instrumento	Precio sin smile	Precio con smile
Caplet ITM 1	1.0644	1.0644
Caplet ITM 2	2.3519	2.3519
Caplet ITM 3	3.2056	3.2056
Caplet OTM 1	0.0263	0.0364
Caplet OTM 2	0.2414	0.2916
Caplet OTM 3	0.54	0.6220
Floorlet OTM 1	0.5197	0.5623
Floorlet OTM 2	0.2236	0.2641
Floorlet OTM 3	0.1151	0.1489
FRN 1	33.0485	33.0485
FRN 2	33.5909	33.5909
FRN 3	33.4308	33.4308

Cuadro 5.2: Precios con y sin *smile* de los instrumentos de referencia.

Aplicando el algoritmo, tras 12 iteraciones hemos obtenido el vector $\mathbf{\Lambda}$ óptimo:

$$\mathbf{\Lambda} = (-1.821, -3.166, -6.315, 0.48, 0.364, 0.394, 1.92, 3.383, 6.995, 1.656, 2.975, 6.091),$$

el cual da el valor mínimo para $W(\mathbf{\Lambda})$, 8.9776. Mediante la comprobación en pantalla, hemos visto que el gradiente de $W(\mathbf{\Lambda})$ iba decreciendo conforme se iba iterando. Finalmente,

el valor del gradiente que hemos aceptado como nulo es $1.4943 \cdot 10^{-13}$. Además, la entropía relativa (mínima) que se alcanza con la probabilidad a posteriori obtenida es 0.2327.

Asimismo, los precios obtenidos al hacer $\mathbb{E}^{\mathcal{P}}[g_j]$ (con \mathbf{p} el vector de probabilidades a posteriori que hemos obtenido) coinciden con los precios con smile de la tabla 5.2. Si en vez de usar \mathbf{p} , calculamos $\mathbb{E}^{\mathcal{Q}}[g_j]$, donde \mathcal{Q} asigna a todas las simulaciones los pesos $1/10000$, obtenemos los precios sin smile recogidos también en la tabla 5.2. Por tanto, el algoritmo ha funcionado correctamente y ya disponemos de las probabilidades para valorar teniendo en cuenta el smile de volatilidad.

5.4. Valoración del swap

El swap, como ya comentamos, es un instrumento que involucra a dos partes, las cuales intercambian pagos de intereses sobre un nominal. En nuestro ejemplo consideramos un swap *payer*, esto es, un swap en el que pagamos el interés fijo y recibimos el variable. En este swap los pagos son semestrales, comienza en $t = 0.5$ y finaliza en $T = 2$. La estructura temporal del swap se puede ver reflejada en las siguientes fechas de fijación y pago:

- Las fechas de fijación son: $t_0 = 0.5$, $t_1 = 1$ y $t_2 = 1.5$.
- Las fechas de pagos son: $t_1 = 1$, $t_2 = 1.5$ y $t_3 = 2$.

Al inicio de cada periodo $[t_{n-1}, t_n]$, con $n = 1, 2, 3$, se fija el tipo Libor para ese periodo de amplitud $\Delta t = t_n - t_{n-1}$, que será el que pague la rama variable en t_n . Así, el pago que hace cada rama en t_n , con $n = 1, 2, 3$, es:

- Pago de la rama variable en $t_n = \mathcal{N} \times \Delta t \times L(t_{n-1}, t_{n-1}, t_n)$.
- Pago de la rama fija en $t_n = \mathcal{N} \times \Delta t \times K$.

En nuestro caso, queremos valorar el swap en una partición de tiempos diaria, con paso temporal $\tau = \left(\frac{1}{30}\right) \left(\frac{1}{12}\right)$, entre el instante 0 y $T = 2$. Cuando tengamos la valoración en todos los instantes de tiempo y para todas las simulaciones, procederemos a obtener las métricas de exposición, como se verá en la siguiente sección. Veamos ahora cómo valorar el swap correctamente.

Para cada tiempo t en el que valoremos el swap, el valor del mismo será la suma de sus pagos futuros descontados hasta tiempo t . En esta suma, si el pago es variable aparecerá con signo positivo y, si es fijo, con signo negativo (pues pagamos la parte fija y recibimos la variable). Recordemos que el Libor visto en t para el periodo $[t_{n-1}, t_n]$ se puede expresar en función de los factores de descuento de la siguiente manera:

$$L(t, t_{n-1}, t_n) = \frac{1}{\Delta t} \left(\frac{P(t, t_{n-1})}{P(t, t_n)} - 1 \right).$$

Teniendo en cuenta todo lo anterior, a tiempo t los pagos en t_n se verán como sigue:

- Pago de la **rama variable** en t_n , visto en $t < t_{n-1}$ (es decir, pago **desconocido**, pues aún no se ha fijado):

$$\begin{aligned}
 \text{Pago variable desconocido} &= P(t, t_n) \times \mathcal{N} \times \Delta t \times L(t_{n-1}, t_{n-1}, t_n) = \\
 &= P(t, t_n) \times \mathcal{N} \times \Delta t \times \frac{1}{\Delta t} \left(\frac{P(t_{n-1}, t_{n-1})}{P(t_{n-1}, t_n)} - 1 \right) = \\
 &= P(t, t_n) \times \mathcal{N} \times \left(\frac{1}{P(t_{n-1}, t_n)} - 1 \right) = \\
 &= P(t, t_n) \times \mathcal{N} \times \left(\frac{1 - P(t_{n-1}, t_n)}{P(t_{n-1}, t_n)} \right) = \\
 &= \mathcal{N} \times \frac{P(t, t_n)}{P(t_{n-1}, t_n)} \times (1 - P(t_{n-1}, t_n)) = \\
 &= \mathcal{N} \times P(t, t_{n-1}) \times (1 - P(t_{n-1}, t_n)) = \\
 &= \mathcal{N} \times (P(t, t_{n-1}) - P(t, t_{n-1}) \times P(t_{n-1}, t_n)) = \\
 &= \mathcal{N} \times (P(t, t_{n-1}) - P(t, t_n)).
 \end{aligned}$$

- Pago de la **rama variable** en t_n , visto en $t \geq t_{n-1}$ (es decir, pago **conocido**, pues ya se ha fijado):

$$\text{Pago variable conocido} = P(t, t_n) \times \mathcal{N} \times \Delta t \times L(t_{n-1}, t_{n-1}, t_n).$$

- Pago de la **rama fija** en t_n , visto en t :

$$\text{Pago fijo} = P(t, t_n) \times \mathcal{N} \times \Delta t \times K.$$

Para llevar a cabo la valoración, en cada tiempo t en el que obtengamos el valor del swap tenemos que seleccionar los pagos que hay pendientes, teniendo en cuenta si estos son conocidos o no. En función de esto, obtendremos el valor del swap en t haciendo la suma de todos los flujos seleccionados anteriormente. Si en el swap pagásemos variable y recibiéramos fijo, entonces el valor del swap en t será como el anterior, pero con signo contrario. En este proceso también tenemos que tener en cuenta que, como criterio a seguir, en el momento en que se realice un pago lo excluirémos del valor total. Es decir, si nos encontramos en una fecha de pago, entonces el valor del swap en esa fecha no llevará incluido el pago que se acaba de realizar.

En la implementación práctica hemos calculado los precios de los bonos mediante $P(t, T) = e^{-\int_t^T r(s)ds}$, aproximando la integral que aparece mediante una suma de Riemann.

Para ilustrar todo lo anterior, veamos cómo son los pagos en algunos instantes temporales del swap que estamos considerando:

- **En $t = t_0$** : Se ha fijado el pago variable que se hará en t_1 , pero no se ha realizado ningún pago todavía. Por ello, aparecerán los 3 pagos pendientes en la parte fija y, en la parte variable, el primer pago será conocido, mientras que el segundo y el tercero no:

- Parte fija:

$$\mathcal{N} \times \Delta t \times K \times (P(t_0, t_1) + P(t_0, t_2) + P(t_0, t_3)).$$

- Parte variable:

$$\mathcal{N} \times (\Delta t \times L(t_0, t_0, t_1) \times P(t_0, t_1) + (P(t_0, t_1) - P(t_0, t_3))).$$

- **En $t_1 < t < t_2$:** En t_1 ya se fijó el pago variable a realizar en t_2 , por lo que en este intervalo de tiempos ese pago ya es conocido. Sin embargo, el pago variable en t_3 aún es desconocido. Por tanto, tenemos:

- Parte fija:

$$\mathcal{N} \times \Delta t \times K \times (P(t, t_2) + P(t, t_3)).$$

- Parte variable:

$$\mathcal{N} \times (\Delta t \times L(t_1, t_1, t_2) \times P(t, t_2) + (P(t, t_2) - P(t, t_3))).$$

- El mismo razonamiento se aplica en el **resto de instantes temporales**.

Una vez valorado el swap en todos los instantes temporales y para todas las simulaciones, hemos representado gráficamente las funciones de distribución y de densidad de las probabilidades a priori y a posteriori. El siguiente gráfico muestra cómo son las funciones de distribución acumuladas:

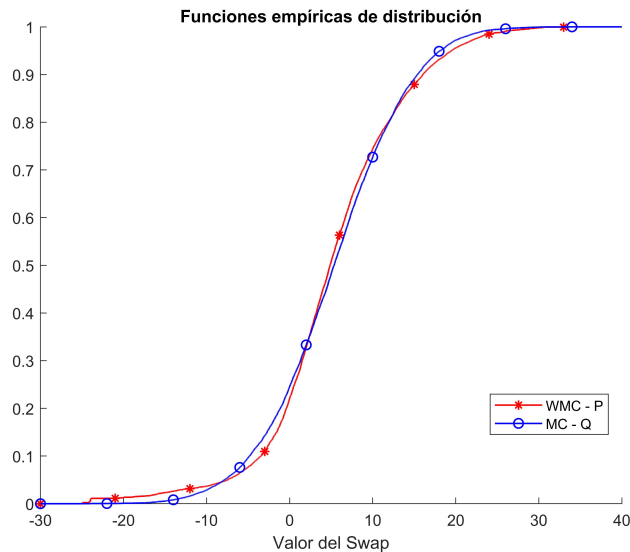


Figura 5.2: Funciones de distribución de las probabilidades a priori y a posteriori.

Observamos que las dos distribuciones son muy parecidas, ya que hemos minimizado su entropía relativa. Sin embargo, si nos fijamos en la zona de la izquierda, vemos que las probabilidades de que el swap tome valores más extremos (que son los que se corresponderán con escenarios más atípicos de los tipos de interés) son más altas con la distribución

a posteriori que con la a priori.

A continuación mostramos las funciones de densidad de las probabilidades a priori y a posteriori:

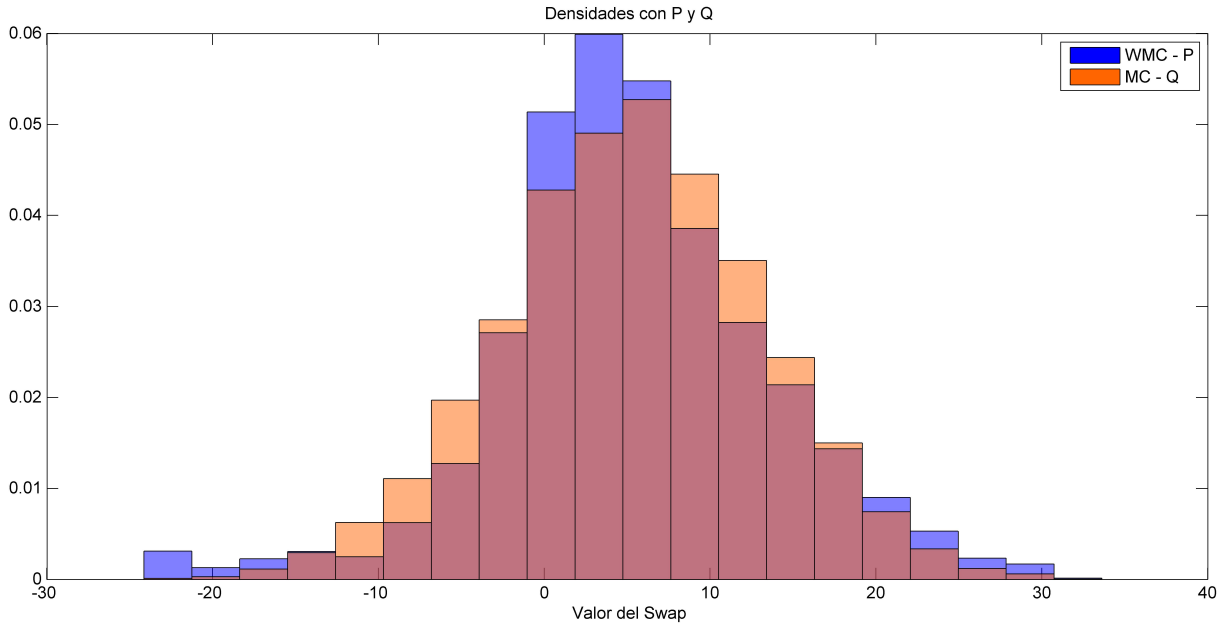


Figura 5.3: Funciones de densidad de las probabilidades a priori y a posteriori.

Como se puede observar, la distribución a posteriori, \mathbf{p} , es más leptocúrtica que \mathbf{q} , es decir, más apuntada en la zona central y con colas más pesadas. Esto se corresponde con la idea que habíamos planteado al comienzo del Trabajo.

5.5. Análisis del riesgo de modelo mediante métricas de exposición

Un objetivo muy importante de este Trabajo es el estudio del riesgo de modelo, en este caso asociado a la no consideración del smile de volatilidad. Para llevar a cabo este análisis, partiremos de las valoraciones del swap que hemos realizado en el apartado anterior. En concreto, hemos valorado el swap a lo largo de una partición de tiempos diaria entre 0 y $T = 2$ en las 10000 simulaciones que teníamos. Esto nos permitirá ahora obtener las métricas de exposición expuestas en el capítulo 4. Como comentamos, estas métricas son estadísticos que cuantifican cuánto nos deberá la contrapartida en el caso de hacer *default* y dependen del instante temporal t en el que las evaluemos. Por este motivo, al realizar la evaluación en toda la partición de tiempos diaria obtendremos perfiles de las mismas.

Para estudiar el riesgo de modelo que se asume al utilizar la probabilidad a priori en vez de la a posteriori como distribución de los valores del swap en cada instante t , vamos a

evaluar las métricas de exposición en dos casos:

- Primero suponiendo que los valores del swap en cada tiempo t se distribuyen según la probabilidad a priori (la que asigna a todas las simulaciones el mismo peso, $1/10000$).
- Después, suponiendo que el swap toma valores en cada tiempo t que están distribuidos según la probabilidad a posteriori, es decir, la que tiene implícita la influencia del smile de volatilidad.

Al hacer esto, para cada una de las métricas de exposición y para los precios MtM obtendremos dos perfiles temporales, uno asociado a la probabilidad a priori y otro a la a posteriori. Veamos qué perfiles hemos obtenido para los precios MtM y para cada métrica, dándoles la interpretación correspondiente:

- **Precios Mark to Market (MtM):**

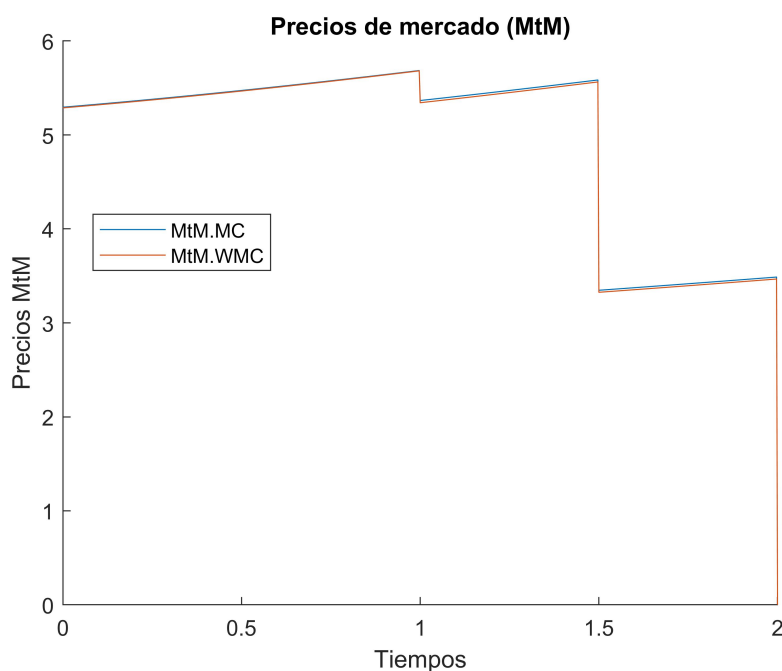


Figura 5.4: Precios Mark to Market del Swap.

Observamos que los precios Mark to Market son muy similares al usar ambas distribuciones. Esto es así porque las dos son medidas de probabilidad martingala, las cuales pueden ser usadas para valorar instrumentos de mercado y, por ello, como suponemos que el mercado no tiene arbitraje, ambas deben dar los mismos precios a los instrumentos. Sin embargo, no son exactamente iguales. Esta pequeña variación se debe a que hemos impuesto la condición de que la probabilidad a posteriori fuera martingala mediante la introducción de 3 restricciones en los precios de los FRN, como comentamos antes. Esta es una condición aproximada y la condición exacta hubiera sido

imponerla para un continuo de vencimientos.

También observamos que el precio del swap, como es esperable, va creciendo ligeramente conforme avanza el tiempo hasta que llega a una fecha de pago. En ese instante, como el pago que se ha producido ya no lo contamos como pendiente, hay un salto negativo en el MtM (que corresponde al pago realizado). Cuando el swap llega a su fin, en el instante $T = 2$, ya tiene valor nulo.

■ **Exposición Positiva Esperada (EPE):**

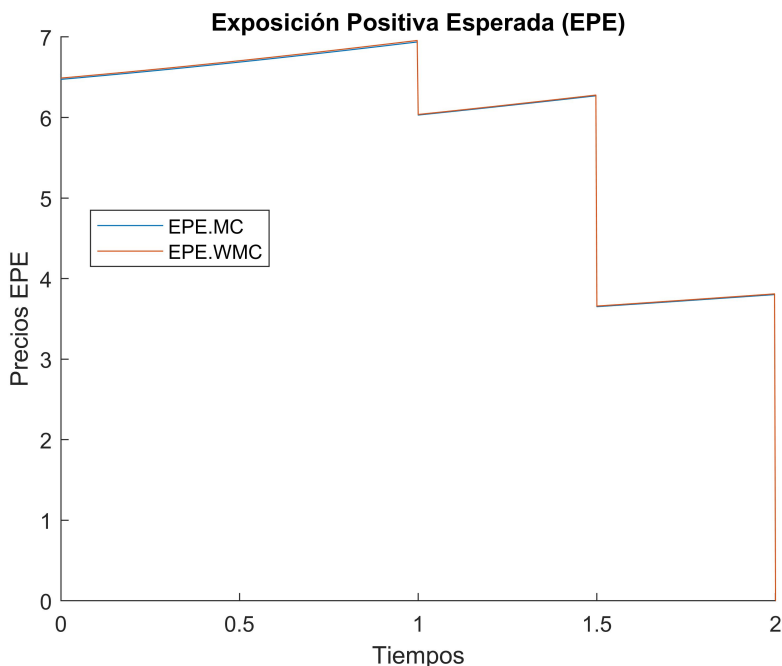


Figura 5.5: Exposición Positiva Esperada del Swap.

La EPE representa cuánto nos pueden deber, en media, en cada instante. De nuevo, hemos obtenido un resultado muy similar con las dos distribuciones. Esto es así porque estas son muy parecidas y, realmente, se está calculando una esperanza en cada una de las medidas de probabilidad. Observamos el mismo comportamiento en cuanto a los pagos que en el Mark to Market. Como la EPE va decreciendo conforme avanza el tiempo, si la contrapartida hiciera *default* en un instante muy avanzado en la vida del swap, la cantidad que nos debería sería menor que si lo hiciera al comienzo.

- **Exposición Negativa Esperada (ENE):**

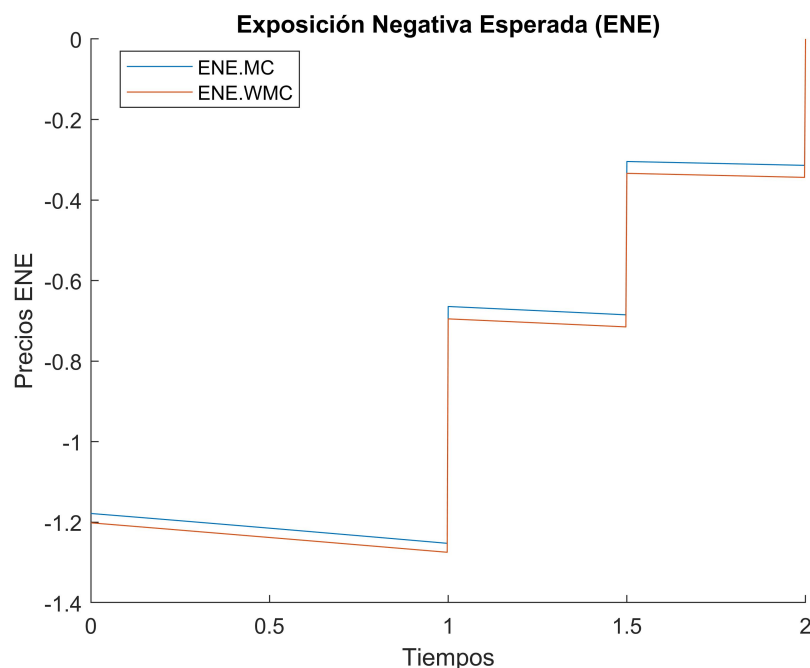


Figura 5.6: Exposición Negativa Esperada del Swap.

La ENE es la EPE que la contrapartida calcula para nosotros. Es decir, es la cantidad, en media, que podemos deber en cada momento. Vemos que también son muy parecidos los resultados obtenidos con las dos distribuciones, aunque no tanto como en los casos anteriores. Observamos que con las probabilidades que incorporan el smile de volatilidad obtenemos una ENE un poco mayor, en el sentido de que podemos deber más a la contrapartida en cada instante que si la calculamos con las probabilidades a priori. Por ello, conviene usar la probabilidad a posteriori ya que es más precisa. Observamos de nuevo que, cuando se realiza un pago, disminuye la cantidad media que podemos deber a la contrapartida, como es de esperar.

- **Exposición Potencial Futura (PFE):**

Como comentamos anteriormente, hemos tomado un nivel de confianza del 95 % para calcular la PFE. Recordemos que esta métrica refleja el precio que se corresponde con el percentil 95 de la distribución de probabilidad que empleemos en cada tiempo. La distribución a posteriori tiene colas más pesadas que la a priori, por lo que, intuitivamente, este hecho debería afectar a los percentiles de ambas distribuciones, propiciando que haya más diferencia en los resultados obtenidos al usar una u otra.

Hemos obtenido los siguientes perfiles:

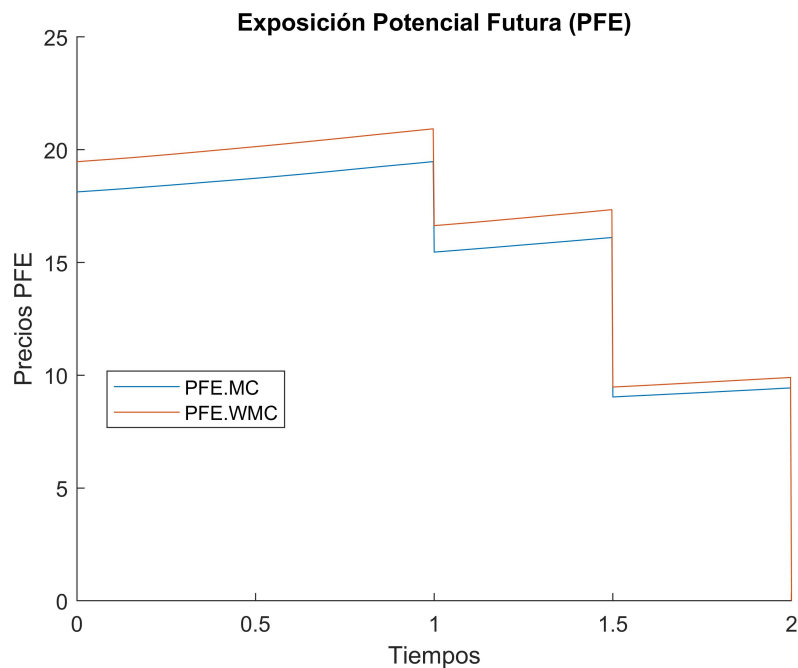


Figura 5.7: Exposición Potencial Futura del Swap.

Como esperábamos, se aprecia diferencia entre la PFE calculada con una y otra distribución. La PFE calculada con la distribución a priori es más baja que la PFE obtenida con la distribución que incluye el smile de volatilidad. Es decir, si no consideramos la influencia del smile, estaremos obteniendo valores para la exposición potencial futura más bajos que los que realmente deberíamos considerar.

La PFE nos indica para cada t un valor de manera que, con una probabilidad del 95 %, el swap valdrá menos que dicho valor. Dicho de otra manera, nos proporciona una cota superior, con una probabilidad del 95 %, para la deuda que se produciría si la contrapartida hiciera default en ese momento. Si usamos la distribución sin smile, esta cota superior será más baja de lo que debería ser, por lo que podríamos tener pérdidas mayores. Aquí se materializa claramente el **riesgo de modelo** asociado a la no consideración del smile.

Capítulo 6

Conclusiones y líneas futuras de investigación

6.1. Conclusiones

El objetivo principal de este Trabajo de Fin de Máster era estudiar y aplicar la metodología de Entropía Relativa Mínima para, partiendo de una probabilidad neutral al riesgo a priori que no considera la existencia del smile de volatilidad, obtener otra más completa que, al valorar bien algunos instrumentos de mercado con smile, incorpore la influencia del mismo. De este modo, es posible objetivar cuánto influye la consideración del smile en las métricas de xVA¹. Así, podemos cuantificar el riesgo de modelo que se deriva de prescindir de la modelización del smile.

Este objetivo es alcanzable mediante la aplicación de la metodología de Entropía Relativa Mínima a la técnica de Monte Carlo. Ésta es conocida como Weighted Monte Carlo o Monte Carlo Ponderado. Particularmente, debido a su naturaleza, esta técnica es especialmente útil en el contexto de xVA, ya que las métricas de exposición siempre son calculadas por Monte Carlo. Uno de los retos a la hora de aplicar la técnica de Monte Carlo Ponderado es garantizar que la probabilidad a posteriori obtenida sea martingala. Para lograrlo, hemos exigido la valoración correcta de los instrumentos conocidos como FRN mediante la inclusión de algunas restricciones más en el problema de minimización. Por ello, una primera conclusión es que la aplicación de la metodología de ERM con instrumentos derivados de los tipos de interés y su aplicación a las métricas de xVA es posible y brinda buenos resultados en la práctica.

El enfoque de Entropía Relativa Mínima tiene varias ventajas con respecto a otras técnicas propuestas para solucionar la problemática del smile de volatilidad. Como se ve en la figura 3.2, podemos descomponer el proceso en varios módulos independientes. De esta manera, tenemos la posibilidad de modificar tanto la distribución de probabilidad a priori como las restricciones en los momentos sin invalidar el resto de la implementación, proporcionándonos una gran flexibilidad. Además, esta técnica posibilita la obtención de

¹xVA (*x - Valuation Adjustments*) es un término para referirse a una serie de ajustes que se hacen en los precios de los derivados. Estos pueden ser de diversa naturaleza, entre los que destacamos el CVA (Credit Valuation Adjustment) y el DVA (Debit Valuation Adjustment). La x denota el tipo de ajuste realizado.

un valor numérico de la entropía relativa que nos da una idea de cuánto se diferencia la distribución a posteriori de la a priori. De esta manera, obtenemos una métrica para cuantificar la influencia del smile de volatilidad. También, al basarse en la simulación, reduce la evaluación computacional de fórmulas grandes y costosas. Por otro lado, como se puede ver en [1], la metodología de ERM tiene implicaciones claras y muy directas para la cobertura de derivados. Aunque esto último no se ha estudiado en este Trabajo, constituye una herramienta muy potente, a la vez que sencilla, a la hora de elaborar carteras de cobertura. Además, nos permite medir el error de cobertura, determinando la bondad de dicha cartera.

Por otro lado, como ha quedado vigente en el análisis de las métricas de exposición para el swap, observamos que es necesario tener en cuenta el smile de volatilidad a la hora de estudiar la exposición. De no hacerlo, incurriremos en un error por infraestimación de la misma que puede tener consecuencias indeseadas cuando se trabaja con carteras grandes. Todo esto, entonces, justifica y favorece la incorporación esta metodología en las instituciones financieras, como un sistema para tener en cuenta la existencia del smile de volatilidad en el mercado.

6.2. Futuras líneas de investigación

Este es un campo totalmente abierto y en expansión. Proponemos algunas de las líneas futuras de investigación que, generalmente, constituyen una extensión de la metodología con la que hemos trabajado:

- Por un lado, la expansión natural de este Trabajo consiste en el desarrollo de la parte de cobertura para instrumentos derivados de los tipos de interés y la inclusión de las métricas de exposición por riesgo de crédito. Esto permitiría, además, el cálculo de sensibilidades (por ejemplo, CVA o DVA) de forma muy eficiente.
- La extensión de los modelos de simulación de los tipos de interés a familias más complejas, e.g., el modelo de Hull-White o el Libor Market Model.
- A medida que se añadan más instrumentos de referencia, tendremos más restricciones sobre la distribución a posteriori. Teniendo en cuenta el coste computacional, se podría investigar cuántos instrumentos de referencia son necesarios para reflejar con alta precisión el smile de volatilidad en dicha medida a posteriori. Esto es así porque podemos ver el problema de la obtención de la distribución a posteriori como un problema de interpolación en los instrumentos de referencia.
- Ligada con la línea de investigación anterior, se podría tratar de determinar de manera objetiva qué instrumentos de mercado conviene incluir para poder obtener la probabilidad a posteriori que mejor refleje la influencia del smile. Aquí habría que considerar los aspectos ya comentados anteriormente que se deben exigir a estos instrumentos de referencia para que el algoritmo converja sin problemas.
- Por último, se puede trabajar en la determinación de una métrica, basada en la distancia de Kullback-Leibler, que permita discernir cuándo las exposiciones potenciales futuras deben considerar el smile o cuándo no es tan necesario. De esa manera, podríamos cuantificar la necesidad de aplicar esta técnica.

Anexo A: Teoremas Fundamentales de Valoración Financiera

En este anexo se presentan los dos Teoremas Fundamentales de Valoración Financiera a los que hacemos referencia en algunos puntos del Trabajo. Antes de enunciarlos, introduciremos dos definiciones que conviene conocer para comprenderlos.

Definición 1 (Oportunidad de arbitraje). *Una oportunidad de arbitraje es la posibilidad de realizar una estrategia financiera V , cuyo precio inicial es negativo y que, sin embargo, a tiempo $T > 0$ tendrá valor no negativo en todos los estados del mercado. Es decir:*

$$V_0 < 0, \quad V_T \geq 0 \quad \text{para todo escenario.}$$

Otra posibilidad también considerada como oportunidad de arbitraje es el caso en el que el valor de la estrategia financiera a tiempo 0 es nulo y a tiempo T es no negativo en todos los estados del mercado, siendo estrictamente positivo en al menos uno de ellos:

$$V_0 = 0, \quad V_T \geq 0, \quad \text{con } V_T > 0 \text{ en algún escenario.}$$

Definición 2 (Mercado completo). *Decimos que el mercado es completo cuando todo instrumento financiero puede ser replicado con activos del mercado.*

Teorema 1 (Primer Teorema Fundamental de Valoración Financiera). *El mercado está libre de oportunidades de arbitraje si y solo si, para cada activo N^2 cuyos precios sean estrictamente positivos, existe una medida de probabilidad tal que el cociente entre cualquier otro activo del mercado, A , y N es un proceso de martingala. Es decir, si para todo $t < T$:*

$$\frac{A_t}{N_t} = \mathbb{E}_t^{\mathbb{Q}} \left[\frac{A_T}{N_T} \right],$$

donde $\mathbb{E}_t^{\mathbb{Q}}$ representa la esperanza en la medida martingala equivalente³ asociada al numerario N y condicionada al estado del mercado a tiempo t .

Teorema 2 (Segundo Teorema Fundamental de Valoración Financiera). *Supongamos que en el mercado existe un activo N cuyos precios son estrictamente positivos y que no hay oportunidades de arbitraje. Entonces, el mercado es completo si y solo si existe una única medida de probabilidad martingala, \mathbb{Q} , asociada a N .*

²Este activo se conoce como *numerario*.

³Si usamos como numerario la cuenta bancaria, es decir, si $N_t = B_t$, entonces la medida martingala equivalente se conoce como **medida neutral al riesgo**.

Bibliografía

- [1] Arrieta, D., 2013. *Entropía relativa y cobertura de derivados*. Ph. D. Thesis, Departamento de Economía Financiera y Contabilidad I, Universidad Complutense de Madrid.
- [2] Arrieta, D., 2015. Minimum Relative Entropy and Cliquet Hedging. *Wilmott*, Vol. 2015, Issue 78, Pages 71-81.
- [3] Avellaneda, M., 1998. Minimum-entropy calibration of asset-pricing models. *International Journal of Theoretical and Applied Finance*, Vol. 1 No. 4.
- [4] Avellaneda, M., Buff, R., Friedman, C., Grandchamp, N., Kruk, L., Newman, J., 2001. Weighted Monte Carlo: a new technique for calibrating asset-pricing models. *International Journal of Theoretical and Applied Finance* Vol. 4, No. 1.
- [5] Cover, T., Thomas, J., 2006. *Elements of Information Theory*. 2nd Edition, Wiley.
- [6] Elices, A., Giménez, E., 2006. Weighted Monte Carlo: Calibrating the Smile and Preserving Martingale Condition. *Risk Magazine*, Vol. 19, No. 5.
- [7] Filipović, D., 2009. Materiales del curso *Interest Rate Models*, organizado por la École Polytechnique Fédérale de Lausanne. URL del curso: <https://es.coursera.org/learn/interest-rate-models>.
- [8] Kullback, S., Leibler, R.A., 1951. On Information and Sufficiency. *Annals of Mathematical Statistics* 22 (1), pp. 7986.
- [9] Oleaga, G., 2018. Apuntes de las asignaturas *Fundamentos de Matemática Financiera* y *Tipos de Interés* del Máster en Ingeniería Matemática de la Universidad Complutense de Madrid.
- [10] Ruiz, I., 2015. *XVA Desks A New Era for Risk Management*. Applied Quantitative Finance series, Palgrave Macmillan.
- [11] Shreve, S.E., 2004. *Stochastic Calculus for Finance II. Continuous-Time Models*. Springer Finance series, Springer.