

Calibración del Modelo de Volatilidad Estocástica de Heston (El Caso de una Opción en Divisas)

MANUEL D. ECHEVERRY

TRABAJO DE TESIS PARA OPTAR AL TÍTULO DE
MAGÍSTER EN CIENCIAS - MATEMÁTICA APLICADA

DIRECTOR
CÉSAR AUGUSTO GOMÉZ, PH.D.



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE MATEMÁTICAS
MEDELLÍN
FEBRERO 2015

Índice general

Índice general	I
Índice de tablas	III
Índice de figuras	IV
Introducción	V
1. Modelo Heston de Volatilidad Estocástica	1
2. Solución de Problemas no Lineales	5
2.1. Métodos de Gauss-Newton y Levenberg-Marquart	6
2.2. Regularización de Tychonoff	7
3. Calibración del Modelo Basada en Minimización de Entropía Relativa	9
3.1. Minimización de Entropía Relativa	10
3.2. Una Aproximación de la Entropía Relativa por medio de Árboles Trinomiales	11
3.3. El Modelo de Heston y la Noción de Entropía Relativa Juntos.	14
4. Enfoque Numérico del Problema de Optimización	16
4.1. Aproximación del modelo de Heston	17
4.2. Implementación del Árbol Recombinante	20
4.3. Optimización de λ y Θ	22
5. Resultados	25
5.1. Árbol Recombinante	25
5.2. Regularización Vía Tychonoff	26
6. Conclusiones	28

A. Definición Intuitiva de la Entropía	29
B. Entropía Relativa en el caso de un solo vencimiento	30
C. Regularidad de la Solución	33
D. Sobre la Aproximación de $\mathbb{H}(\mathbb{Q}; \mathbb{P})$ en un Espacio de Caminos	35
Bibliografía	37

Índice de tablas

5.1. Resultados del proceso de calibración con λ variable y con 20 pasos temporales.	25
5.2. Resultados del proceso de calibración con 40 pasos temporales.	25
5.3. Precios de opciones y strikes que se usaron en la calibración.	26
5.4. Resultados del experimento.	27

Índice de figuras

2.1. L-curva creada a partir de un ejemplo de <i>Regularization Tools, A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems</i> , Per Christian Hansen, 2008.	8
4.1. Marchas discretas.	19

Introducción

Hoy en día, el modelamiento matemático de mercados financieros es un tópico de interés en el ámbito de las matemáticas aplicadas, la estadística, la economía y principalmente en la industria. En general, los modelos están buscando simular algunos instrumentos financieros en aras de aumentar su concordancia con la realidad. En este trabajo se abordarán varios aspectos del modelo de volatilidad estocástica de Heston para la valoración de opciones el cual extiende el celebre modelo de Black-Scholes.

Imagine que usted pretende hacer una inversión en yuanes¹ y pretende que pasado un tiempo pueda recibir ganancias por las variaciones de la tasa de cambio, este es un negocio riesgoso, ya que la tasa podría variar demasiado en su contra y usted perdería mucho dinero. Puede pensar que sería bueno tener una forma de estar seguro de que lo que reciba el día que pretende hacer el cambio no le sea demasiado desfavorable, y hay una entidad que le ofrece asumir el riesgo de comprometerse a comprarle sus yuanes si usted lo desea a una tasa preestablecida, pero dejándole con la libertad de vendérselos a otros si quiere. Si lo ve de esta forma usted poseería un producto financiero al que llamaremos una opción de venta, algo similar quisiera alguien que en el futuro desea comprar algo y quiere que sea a cierto precio, lo cual sería una opción de compra. Las opciones son un tipo de derivado financiero debido a que su valor se deriva del valor de otro activo denominado subyacente, en el ejemplo anterior sería la tasa de cambio entre yuanes y dólares.

Una opción europea² es un contrato entre dos entidades, el emisor y el tenedor, en el que el tenedor adquiere el derecho por medio del pago de una prima al emisor de comprar un activo en una fecha futura T por un valor que es preestablecido y que se denomina precio de ejercicio K . Esto puede verse como un instrumento que permite fijar los precios en el futuro o como un instrumento de cobertura.

El emisor mismo adquiere el deber de vender al tenedor dicho activo al precio de ejercicio de la opción. El nombre de este contrato se le da por el hecho de que el tenedor en el caso de que el valor al vencimiento del activo S_T sea menor que K entonces este podría ejercer el derecho de comprar al valor preestablecido. Lo anterior se formula así:

$$(S_T - K)^+ := \max\{0, S_T - k\}$$

¹Moneda de China.

²Existen otros tipos de opciones, como la americana y la asiática.

La opción que hemos descrito se denomina opción call o de compra, y una en la cual el tenedor tiene el derecho de vender el bien por un precio de ejercicio, se llama de venta o put. El derecho de compra o venta en la opción descrita arriba sólo puede ejercerse en el momento T y se le da el nombre de opción europea, en el caso de que en una opción se pueda hacer uso del derecho en cualquier momento hasta T se denomina de tipo americano (opción americana). Nos ocuparemos de opciones en divisas en las cuales S_t modela la evolución de una tasa de cambio.

Habiéndose vendido la opción, el emisor usa la prima para realizar una proceso de cobertura (Hedging) destinado a reducir al mínimo el riesgo de incumplimiento del mismo hacia el tenedor en caso de que este ejerza la opción. El proceso de cobertura es en sí mismo uno de los propósitos que le da importancia a la estimación de los parámetros que otorga la calibración, ya que el cubrimiento requiere encontrar el valor de la derivada

$$\frac{\partial}{\partial S} C_{\text{Heston}}(S_t, v_t, \bar{\Theta})$$

donde $\bar{\Theta}$ es el vector que contiene los parámetros hallados por medio de la calibración.

La calibración consiste en estimar los valores de los parámetros del modelo de tal forma que los precios de las opciones deducidos de este reproduzcan o aproximen los valores de las opciones observados en el mercado. Estos corresponden a precios obtenidos por la especulación y las tendencias del mercado, justamente el mercado que pretende ser simulado por el modelo que vamos a calibrar. Nuestro propósito principal será presentar una calibración basada en la minimización de entropía relativa³ para luego compararla con estimación basada en el método de Tychonoff⁴.

³Este término será descrito más adelante.

⁴Será definido posteriormente.

Modelo Heston de Volatilidad Estocástica

Antes de presentar la idea que lleva al establecimiento del modelo, presentaremos algunas definiciones importantes, las cuales son presentadas en la referencia [11] en sus dos primeros capítulos.

Definición 1.1. Dado un conjunto Ω y \mathcal{F} una σ -álgebra de subconjuntos de Ω , definimos una filtración $\{\mathcal{F}_t\}_{t \geq 0}$ como una familia de sub- σ -álgebras de \mathcal{F} tales que, $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$, si $0 \leq s \leq t \leq \infty$.

Definición 1.2. Dado un proceso estocástico X definimos la filtración generada por el como

$$\mathcal{F}_t^X := \sigma(X_s; 0 \leq s \leq t)$$

Definición 1.3. Un proceso estocástico X es adaptado a la filtración $\{\mathcal{F}_t\}$ si, para cada $t \geq 0$, X_t es una variable aleatoria \mathcal{F}_t -medible.

Definición 1.4. Un movimiento browniano estándar (unidimensional) es un proceso de trayectorias continuas y adaptado $B_t = \{B_t, \mathcal{F}_t; 0 \leq t < \infty\}$, definido en algún espacio de probabilidad (Ω, \mathcal{F}, P) , con las siguientes propiedades

- (i) $B_0 = 0$.
- (ii) Para $0 \leq s < t$ el incremento $B_t - B_s$ es independiente de \mathcal{F}_s y es normalmente distribuido con media cero y varianza $t - s$.

Habiendo definido el movimiento browniano, comenzaremos a introducir el modelo. En los años setenta Fisher Black y Myron Scholes propusieron un modelo para la valoración de opciones que consta de una ecuación diferencial estocástica que tenía en cuenta la volatilidad del precio de un activo y la tasa libre de riesgo del mismo, esta volatilidad era considerada constante y por lo tanto al hacerla variable se podría extender de manera natural a un modelo más complejo al permitir que dicha volatilidad sea aleatoria, tal como lo hizo Steven L. Heston en su celebre modelo que trataremos a continuación.

El Modelo Heston describe el cambio en el precio de un bien, asumiendo que la volatilidad es variable y estocástica. Ambos, el precio del activo y la volatilidad, son procesos estocásticos que poseen ecuaciones similares en la que el primero depende del segundo.

En el modelo de Black-Scholes se asume que la variación en el precio de un activo S en un periodo corto de tiempo dt será igual al precio en el instante t por su tasa de crecimiento (o decrecimiento) más el cambio generado por la volatilidad el cual se modela por medio de suponerla una constante \sqrt{v} multiplicada por el incremento de un movimiento browniano W_1

$$\begin{aligned} dS_t &= S_t (r dt + \sqrt{v} dW_1) \\ &= r S_t dt + S_t \sqrt{v} dW_1 \end{aligned} \quad (1.1)$$

En el modelo de volatilidad estocástica, la volatilidad sigue un proceso estocástico parecido y agregamos una ecuación extra al sistema que corresponde a la correlación entre los dos movimientos brownianos que tiene el sistema, en concreto, el modelo Heston se presenta de la siguiente forma para una tasa de cambio S_t , o sea, se representa el precio de una divisa en términos de otra:

$$\begin{aligned} dS_t &= (\bar{r} - r_f) S_t dt + \sqrt{v_t} S_t dW_1 \\ dv_t &= \kappa(\theta - v_t) dt + \sigma \sqrt{v_t} dW_2 \end{aligned} \quad (1.2)$$

donde W_1 y W_2 son movimientos brownianos correlacionados, lo cual quiere decir que cuando uno de los procesos tiene un incremento la otra lo tendrá y viceversa, esto se expresa en la siguiente ecuación

$$\mathbb{E} [dW_1 \cdot dW_2] = \rho dt$$

además

- \bar{r} es la tasa libre de riesgo doméstica, en nuestro caso nos referimos a la tasa de interés de un bono del tesoro del país que emite la moneda que tomamos como base para representar a la otra divisa.
- r_f es la tasa libre de riesgo extranjera, sería la tasa de cambio de un bono del tesoro del estado emisor de la divisa cuyo valor es expresado en términos de la otra¹
- v_t es la volatilidad del activo subyacente S_t en el momento t
- σ es la *volatilidad de la volatilidad*
- κ es la tasa de reversión a la media, es una medida de la rapidez con la cual la volatilidad revierte a su media a largo plazo θ .
- θ es la media de la volatilidad a largo plazo.
- $W_{1,2}$ son movimientos brownianos.
- ρ es el parámetro de correlación entre W_1 y W_2 .

Observación 1.1. *Suponemos que el sistema 1.2 está escrito con respecto a una medida martingala equivalente, esto puede ser visto en la referencia [9] sección 1.4.*

¹En nuestro caso particular trabajamos con la tasa de cambio entre euro y dólar (EUR/USD) entonces nuestra tasa libre de riesgo doméstica la definida por la FED (Reserva Federal) en Estados Unidos y la tasa libre de riesgo extranjera será impuesta por el Banco Central Europeo.

Observación 1.2. *En el contexto del modelo de Heston, como se sabe, no hay una única medida martingala equivalente, pero suponemos que las ecuaciones 1.2 están escritas respecto a la medida martingala “seleccionada por el mercado, en el sentido de que bajo esta medida los precios observados de las opciones se ajustan a los precios teóricos proporcionados por el modelo de Heston”. Esto puede ser consultado en la referencia [9] sección 2.5.*

Observación 1.3. *Para asegurar que el proceso v_t sea estrictamente positivo de darse que $2\theta\kappa > \sigma^2$, esta es conocida como la condición de Feller.*

Ahora, el precio de una opción basada en este modelo viene dada por

$$C(t, S, v; T) = E \left[e^{-r(T-t)} (S_T - K)^+ | S_t = S, v_t = v \right]$$

donde t es el momento actual, T es el vencimiento, S el subyacente y v su volatilidad. Simbolizaremos su valor con C y puede demostrarse que la función anterior satisface la siguiente ecuación diferencial parcial, como se muestra en la referencia [14]

$$\begin{aligned} \frac{\partial C}{\partial t} + \frac{S^2 v}{2} \frac{\partial^2 C}{\partial S^2} + (\bar{r} - r_f) S \frac{\partial C}{\partial S} - (\bar{r} - r_f) C \\ + [\kappa(\theta - v) - \lambda v] \frac{\partial C}{\partial v} + \frac{\sigma^2 v}{2} \frac{\partial^2 C}{\partial v^2} + \rho \sigma S v \frac{\partial^2 C}{\partial S \partial v} = 0 \end{aligned}$$

una solución semi-explicita con condición final en tiempo T , $C(T, S, v; T) = (S_T - K)^+$, de esta ecuación fue desarrollada por Heston vía el método de funciones características en la referencia [10], dicha solución es

$$\begin{aligned} C(t, S_0, v_0; T) &= S P_1 - K e^{-(\bar{r} - r_f)(T-t)} P_2 \\ P_j &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[\frac{e^{-iu \ln(K)} \Phi_j(S_0, V_0, t, T, u)}{iu} \right] du, j = 1, 2 \end{aligned}$$

Finalmente, la solución es encontrada haciendo uso de las fórmulas anteriores y de la transformada de Fourier para calcular las funciones P_1 y P_2

$$\begin{aligned} \Phi_j(S_0, V_0, t, T, u_j) &= \exp\{C_j(\tau; \phi) - D_j(\tau; \phi)V_0 + i\phi S_0\}, \tau = T - t \\ C_j(\tau, \phi) &= (\bar{r} - r_f)\phi i\tau + \\ &\quad \frac{\kappa\theta}{\sigma^2} \left[(b_j - \rho\sigma\phi i + d_j)\tau - 2\ln \left(\frac{b_j - \rho\sigma\phi i + d_j}{1 - g_j} \right) \right] \\ D_j(\tau, \phi) &= \frac{b_j - \rho\sigma\phi i + d_j}{\sigma^2} \left(\frac{1 - e^{d\tau}}{1 - g e^{d\tau}} \right) \end{aligned}$$

donde,

$$\begin{aligned} g_j &= \frac{b_j - \rho\sigma\phi i + d_j}{b_j - \rho\sigma\phi i - d_j} \\ d_j &= \sqrt{(\rho\sigma\phi i - b_j)^2 - \sigma^2(2u_j\phi i - \phi^2)} \\ u_1 &= 0,5 \quad , \quad u_2 = -0,5, \\ b_1 &= \kappa + \lambda - \rho\sigma \quad , \quad b_2 = \kappa + \lambda \end{aligned}$$

Nuestro objetivo es calibrar los parámetros κ , θ , σ y ρ del sistema 1.2, i.e. de forma tal que se minimice la suma de los cuadrados de las diferencias entre los valores encontrados vía una función que calcule el precio de la opción y los valores de opciones observados en el mercado financiero. Tenemos esencialmente un problema inverso, para más detalles sobre este tema véase la referencia [4].

De forma similar como se modificó el modelo Black-Scholes por medio de suponer una volatilidad estocástica, el modelo Heston también puede modificarse asumiendo que algunos de los parámetros son dependientes del tiempo, esto es expuesto en la referencia [15] y ahí es presentado con un alto grado de generalidad

$$\begin{aligned}dX_t &= \sqrt{v_t}S_t dW_1 - \frac{v_t}{2}dt, X_0 = x_0 \\dv_t &= \kappa(\theta_t - v_t)dt + \sigma_t\sqrt{v_t}dW_2 \\d(W_1, W_2)_t &= \rho_t dt\end{aligned}\tag{1.3}$$

Este modelo no será tratado explícitamente en este trabajo, sin embargo creemos que la técnica de calibración que estudiamos puede ser también adoptada para este modelo.

CAPÍTULO 2

Solución de Problemas no Lineales

El problema que tratamos de solucionar es evidentemente no lineal, afortunadamente existen diversos métodos para dar solución a problemas del tipo que estamos trabajando. En este capítulo daré una breve introducción a algunos de estos métodos.

Supongamos que se tiene una ecuación de la forma $F(x) = 0$ con n ecuaciones y n incógnitas, en caso de que la función F sea diferenciable se puede aproximar el valor de la función a través de su expansión en serie de Taylor al rededor de un punto, digamos x_0

$$F(x) \approx F(x_0) + J(F(x_0))(x - x_0)$$

donde $J(F(x))$ es el jacobiano.

Si x^* es la solución de nuestro problema y usando la notación $\Delta x = (x^* - x_0)$

$$F(x^*) = 0 \approx F(x_0) + J(F(x_0))\Delta x$$

de lo cual

$$J(x_0)\Delta x = -F(x_0) \tag{2.1}$$

Resolviendo este sistema lineal obtenemos a Δx para mejorar la estimación de x^* . Este proceso se puede repetir hasta obtener un x^* que se acerque a ser una solución del problema. El método que acabamos de describir es conocido como método de Newton, un algoritmo en pseudo-código para el será

Paso 1 Elija un valor inicial para x_0 .

Paso 2 Calcule $J(F(x_0))$ y $F(x_0)$.

Paso 3 Solucione $J(F(x_0))\Delta x = -F(x_0)$

Paso 4 Haga $x^* = \Delta x + x_0$.

Paso 5 Si x^* se aproxima adecuadamente a la solución de nuestro problema, pare.

Si no es así, entonces haga $x_0 = x^*$ y vuelva al paso 2.

Ahora, tenemos una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y suponemos que esta es C^2 , si quisiéramos minimizarla entonces, en caso de que x^* fuera el valor que llega al mínimo, debería pasar

que $\nabla f(x^*) = 0$. Dado que la función es C^2 podemos aproximar al gradiente de la función con una expansión en series de Taylor

$$\begin{aligned}\nabla f(x^*) = 0 &\approx \nabla f(x_0) + H(f(x_0))\Delta x \quad \text{y por lo tanto} \\ H(f(x_0))\Delta x &= \nabla f(x_0)\end{aligned}\tag{2.2}$$

donde $H(f(\cdot))$ es la matriz hessiana.

2.1. Métodos de Gauss-Newton y Levenberg-Marquart

Dado el sistema de ecuaciones de k ecuaciones no lineales $G(y) = d$, puede ser a veces imposible resolverlo de forma exacta, por ejemplo cuando un vector particular de datos d no pertenece al rango de la función G . En este caso podemos optar por buscar una solución tal que minimice

$$f(y) := \sum_{i=1}^k (G(y)_i - d_i)^2\tag{2.3}$$

donde $G(y)_i$ es la i -ésima componente del vector $G(y)$.

Definimos las funciones escalares $f_i(y) = G(y)_i - d_i$ con $i = 1, 2, \dots, k$ y la función

$$F(y) = \begin{bmatrix} f_1(y) \\ \vdots \\ f_k(y) \end{bmatrix}$$

y notemos que $f(y) = \sum_{i=1}^k (f_i(y))^2 = \|F(y)\|_2^2$. Recordemos que buscamos minimizar la distancia entre $G(y)$ y d , lo cual es equivalente a minimizar la función $f(y)$. Como vimos anteriormente, podemos hacer esto si encontramos x tal que $\nabla f(x) = 0$.

Usando las propiedades de la función f llegamos a que su gradiente puede ser escrito como

$$\nabla f(y) = 2J(f(y))^T F(y)\tag{2.4}$$

y haciendo uso de la ecuación 2.2

$$H(f(y))\Delta y = 2J(f(y))^T F(y)\tag{2.5}$$

La matriz hessiana en el caso de f puede ser aproximada calculando la matriz jacobiana

$$H(f(y)) \approx 2J(f(y))^T J(f(y))\tag{2.6}$$

como es descrito detalladamente en [13], y haciendo uso de esta ecuación y la anterior a esta

$$J(f(y))^T J(f(y))\Delta y = J(f(y))^T F(y)\tag{2.7}$$

resolviendo la ecuación anterior para Δy se puede hallar una aproximación a la solución del problema de mínimos cuadrados, este método es llamado el método de Gauss-Newton (GM), y a partir de este se produce una sucesión de vectores que encuentra una solución.

Un problema asociado al método que acabamos de describir es que posiblemente no se llegue a una solución debido a que la matriz $J(f(y))^T J(f(y))$ este cerca o sea una matriz singular, para solucionar este inconveniente el método de Levenberg-Marquart (LM) propone modificar el problema por uno similar pero con posiblemente mejor comportamiento numérico, así, este presenta la siguiente forma

$$(J(f(y))^T J(f(y)) + \lambda I) \Delta y = J(f(y))^T F(y) \quad (2.8)$$

siendo λ un número positivo, la escogencia del mismo varía dependiendo del comportamiento que esté presentando la búsqueda de la solución, no se profundizará mucho en este tema en particular pues nuestro problema será resuelto con otros métodos.

2.2. Regularización de Tychonoff

Una forma alternativa de tratar el problema de mínimos cuadrados no lineales es la regularización de Tychonoff, la cual expondremos en esta sección.

Observación 2.1. *También podría ser importante trabajar con el error relativo, sin embargo hay varios trabajos que consideran el error cuadrático, como en la referencia [14]. En el contexto del modelo de Heston no es clara la adaptación del concepto de volatilidad implicada.*

Sea $\Theta = (\kappa, \theta, \sigma, \rho)$ un vector en \mathbb{R}^4 que minimiza la siguiente expresión

$$\sum_{(K,T) \in \mathcal{K}} \|C(K, T; \Theta) - C^{obs}(K, T)\|^2 \quad (2.9)$$

donde \mathcal{K} es un conjunto de pares (K, T) de precio de ejercicio y vencimiento. La componente $C(K : T; \Theta)$ nos da el valor de la opción en función en Θ y $C^{obs}(K, T)$ es el valor de una opción tomado del mercado para el strike (precio de ejercicio) y vencimiento correspondientes. El método de Tychonoff consiste en considerar soluciones Θ de 2.9 cumplen $\|\Theta - \Theta_0\|^2 < \delta$, donde Θ_0 representa una solución a priori de los parámetros, por ejemplo el resultado de una calibración anterior. Solucionaríamos entonces:

$$\begin{cases} \min_{\Theta} \sum_{(K,T) \in \mathcal{K}} \|C(K, T; \Theta) - C^{obs}(K, T)\|^2 \\ \|\Theta - \Theta_0\| < \delta \end{cases} \quad (2.10)$$

También se puede pensar en minimizar la norma de $\Theta - \Theta_0$ y que 2.9 no se salga de una tolerancia $\epsilon > 0$

$$\begin{cases} \min \|\Theta - \Theta_0\| \\ \sum_{(K,T) \in \mathcal{K}} \|C(K, T; \Theta) - C^{obs}(K, T)\|^2 < \epsilon \end{cases} \quad (2.11)$$

Se puede pensar en un parámetro que penalice no sólo las soluciones grandes sino también las pequeñas, dependiendo del caso, y así no separar el problema en una minimización acompañada de una restricción. Esta idea se plantearía con una optimización de la siguiente

forma:

$$\min_{\Theta} \left\{ \sum_{(K,T) \in \mathcal{K}} \|C(K, T; \Theta) - C^{obs}(K, T)\|^2 + \alpha \|\Theta - \Theta_0\|^2 \right\} \quad (2.12)$$

Observación 2.2. Las ecuaciones 2.10, 2.11 y 2.12 son equivalentes en el sentido de que una adecuada escogencia de δ , ϵ y α nos llevaría a encontrar la misma solución, tal como se menciona en la referencia [13] página 94.

La escogencia de α es un tema bastante estudiado en el caso en el que el problema a minimizar es una función lineal en los parámetros, pero para nuestro caso, al no serlo, el problema se convierte más en ensayo y error que otra cosa, aunque existe un método que se usa para el caso lineal que podría funcionar para nuestro fin, la L-curva. Para mayor profundidad sobre este método se recomienda la referencia [4].

La L-curva es un método gráfico para elegir el parámetro de regularización, este consiste en graficar bajo una escala $\log - \log$ a $\|\Theta - \Theta_0\|$ versus $\|C(S_0, v_0, t_0; \Theta) - C^{mk}\|$ y tomar a α como el punto de máxima curvatura, este método recibe tal nombre por la forma que tiene al ser graficada.

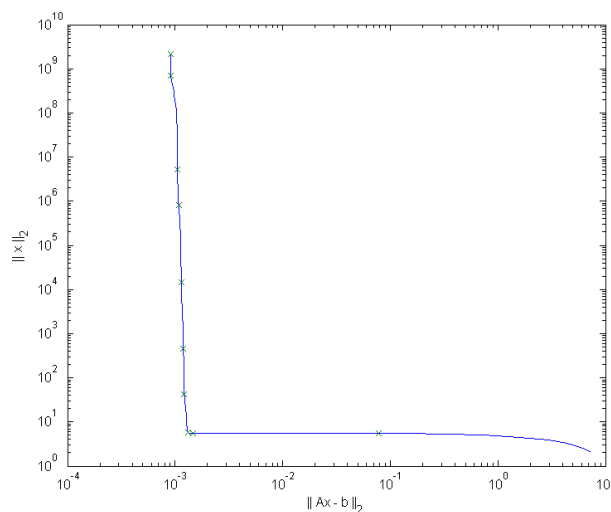


FIGURA 2.1. L-curva creada a partir de un ejemplo de *Regularization Tools, A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems*, Per Christian Hansen, 2008.

La función C calculada usando los códigos propuestos en la referencia [6]. Produciremos vectores aleatorios y haremos experimentos repetidamente tratando de ver si podemos acercarnos a implementar una L-curva para nuestro problema.

También se podría escoger un grupo de parámetros α y comparar las soluciones obtenidas vía una implementación de ASA (Temple Simulado Adaptativo), para así, tener una idea de cuál sería el mejor para nuestra calibración. Los resultados de estos acercamientos serán presentados en una sección posterior.

Calibración del Modelo Basada en Minimización de Entropía Relativa

Definición 3.1. Sean f y f_0 funciones de densidad de probabilidad, definimos la entropía relativa de f respecto a f_0 como:

$$\begin{aligned} \mathbb{H}(f|f_0) &= \mathbb{E}_f \left[\log \left(\frac{f}{f_0} \right) \right] \\ &= \int f \log \left(\frac{f}{f_0} \right) dx \end{aligned}$$

Hay varios nombres y notaciones para el concepto de entropía relativa, algunos nombres de uso común para esta cantidad son número de *Kullback-Leibler*, *divergencia de información*, *Información cruzada*, entre otros.

Usamos la notación \mathbb{E}_f para expresar que el valor esperado que se calcula es en base a la medida de probabilidad que tiene a f como su función de densidad. Observe además que, $\mathbb{H}(f|f_0) = 0$ si $f = f_0$. Más aún, se tiene que

- $\mathbb{H}(f|f_0) \geq 0$ con igualdad si y sólo si $f = f_0$
- $\mathbb{H}(f|f_0)$ es convexa en el par (f, f_0)

Esto último es tomado de la referencia [12] (pags 23,29).

La no negatividad de la entropía relativa nos permite comprender que la entropía relativa representa una medida de “similitud” o una “distancia” entre 2 medidas de probabilidad, en otras palabras la entropía relativa es una medida de que tan diferentes son 2 medidas de probabilidad. Sin embargo la entropía relativa no es simétrica ni satisface la desigualdad triangular por lo que deja de ser una métrica en el sentido matemático estricto. Sin embargo ha sido empleada en varias áreas para comparar distribuciones de probabilidad. La noción de entropía relativa ha sido usada ya en la calibración de modelos financieros, tal como se hizo en la referencia [?].

3.1. Minimización de Entropía Relativa

En esta sección seguiremos de cerca la referencia [2] donde se podrán revisar todos los detalles técnicos.

Considere dos medidas de probabilidad \mathbb{P} y \mathbb{Q} definidas en un espacio de caminos continuos $\Omega = \{S_t, 0 \leq t \leq T\}$ tales que

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu_t^{\mathbb{P}} dt + \sigma_t^{\mathbb{P}} dZ_t^{\mathbb{P}} \text{ bajo } \mathbb{P} \\ \frac{dS_t}{S_t} &= \mu_t^{\mathbb{Q}} dt + \sigma_t^{\mathbb{Q}} dZ_t^{\mathbb{Q}} \text{ bajo } \mathbb{Q} \end{aligned} \tag{3.1}$$

donde $Z_t^{\mathbb{P}}$ y $Z_t^{\mathbb{Q}}$ son movimientos brownianos estándar con respecto a las medidas de probabilidad \mathbb{P} y \mathbb{Q} respectivamente, y μ_i y σ_i son procesos acotados y adaptados a las filtraciones generadas por Z_i , puede verse [11] pp 4-5.

Definición 3.2. Dadas \mathbb{Q} y \mathbb{P} medidas de probabilidad en un espacio de trayectorias, definimos la entropía relativa de \mathbb{Q} respecto a \mathbb{P} como

$$\mathbb{H}(\mathbb{Q}; \mathbb{P}) = \int_{\Omega} \ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) d\mathbb{Q}$$

Observación 3.1. Las definiciones 3.1 y 3.2 definen conceptos distintos, la primera presenta la entropía relativa de una función de densidad de probabilidad con respecto a otra, en cambio 3.2 trata la entropía relativa de una **medida** de probabilidad con respecto a otra donde ambas medidas pueden ser medidas de probabilidad en un mismo espacio de de dimensión infinita. Ahora, $\frac{d\mathbb{Q}}{d\mathbb{P}}$ es la derivada de Radon-Nikodym de \mathbb{Q} con respecto de \mathbb{P} , para mayor profundidad sobre este tópico puede verse [16].

Vamos a considerar el siguiente problema de minimización con restricciones:

$$\inf_{\mathbb{Q}} \mathbb{H}(\mathbb{Q}; \mathbb{P}) \tag{3.2}$$

sujeta a las restricciones

$$e^{-r(T_i)} \mathbb{E}^{\mathbb{Q}}[(S_{T_i} - K_i)^+ | S_{t_0}, v_{t_0}] = C_i^{\text{Obs}}, \quad i = 1, 2, \dots, M. \tag{3.3}$$

Donde $C_i^{\text{Obs}} := C^{\text{Obs}}(T_i, K_i)$ constituye un conjunto de precios observados en el mercado de opciones para determinados pares vencimiento-precio de ejercicio (T_i, K_i) y suponemos que \mathbb{Q} es la medida neutra al riesgo determinada por dicho mercado, pues es la medida de probabilidad utilizada para valorar las opciones bajo consideración.

Las restricciones en (3.3) expresan el hecho de que bajo las medidas de probabilidad consideradas \mathbb{Q} , idealmente los precios de opciones observados C_i^{Obs} sean reproducidos por los precios teóricos según el modelo de Heston

$$\begin{aligned} C_{\text{Heston}}^{\mathbb{Q}}(S_{t_0, v_{t_0}}, t_0; T_i, K_i) &= e^{-r(T_i - t_0)} \mathbb{E}^{\mathbb{Q}}[(S_{T_i} - K_i)^+ | S_{t_0}, v_{t_0}] \\ &= e^{-r(T_i - t_0)} \int_{-\infty}^{\infty} (s - K_i)^+ dF_{S_{T_i}}^{\mathbb{Q}}(s). \end{aligned}$$

Donde la función $F_{S_{T_i}}^{\mathbb{Q}}(s)$ es la función de distribución de S_{T_i} bajo la medida \mathbb{Q} . O sea

$$F_{S_{T_i}}^{\mathbb{Q}}(s) = \mathbb{Q}(S_{T_i} \leq s | S_{t_0}, v_{t_0}).$$

El problema de optimización (3.2) con las restricciones (3.3) será tratado usando el método de multiplicadores de Lagrange, por lo tanto en lo sucesivo vamos a considerar el siguiente problema de optimización:

$$\inf_{\lambda} \sup_{\mathbb{Q}} \left\{ -\mathbb{H}(\mathbb{Q} | \mathbb{P}) + \sum_{i=1}^M \lambda_i \left(e^{-rT_i} \int G_i(s) dF_{T_i}^{\mathbb{Q}}(s) - C_i^{\text{Obs}} \right) \right\} \quad (3.4)$$

donde las funciones $G_i(s)$ determinan los pagos al vencimiento de las correspondientes opciones, es decir $G_i(s) = (s - K_i)^+$.

Observación 3.2. \mathbb{P} es una medida de probabilidad que recoge alguna información econométrica a priori sobre el proceso seguido por la tasa de cambio S_t . Como por ejemplo, estimaciones por medios econométricos de la distribución empírica de la tasa de cambio o también \mathbb{P} puede resumir el resultado de una calibración anterior.

3.2. Una Aproximación de la Entropía Relativa por medio de Árboles Trinomiales

Acá seguimos el planteamiento en [1] de aproximar la entropía relativa

$$\mathbb{H}(\mathbb{Q}; \mathbb{P}) = \int_{\Omega} \ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) d\mathbb{Q}$$

entre las medidas de probabilidad asociadas a los procesos en (3.1) por medio de entropías calculadas sobre un árbol trinomial. Vamos a considerar aproximaciones en tiempo discreto de los procesos en (3.1) y veremos el comportamiento de la sucesión de entropías en el límite cuando $\Delta t \rightarrow 0$.

Sea $\{S_0, S_1, \dots, S_N\}$ una discretización de un camino¹ S_t , $0 \leq t \leq N$. La probabilidad de que el camino $\{S_i\}_{i=0}^N$ ocurra puede ser expresada como

$$\prod_{n=0}^{N-1} \pi_n^{\mathbb{P}}$$

¹Ver Apéndice G.

donde $\pi_n^{\mathbb{P}}$ es la probabilidad condicional con respecto a la medida de probabilidad \mathbb{P} de que el camino tome el valor S_{n+1} en el instante $n + 1$ condicional en que tomó el valor S_n en el instante n .² $\pi_0^{\mathbb{P}}$ sigue la regla general, es decir, $\pi_0^{\mathbb{P}}$ es la probabilidad condicional de que el camino tome el valor S_1 en el instante 1 dado que el camino comienza en S_0 .

De la misma forma $\prod_{n=1}^{N-1} \pi_n^{\mathbb{Q}}$ es la probabilidad de que dicho camino ocurra con respecto a la probabilidad \mathbb{Q} .

Aproximamos entonces la entropía relativa siguiendo[1]

$$\begin{aligned}
 \mathbb{H}(\mathbb{Q}; \mathbb{P}) &= \sum_{\Omega} \ln \left(\frac{\prod_{n=0}^{N-1} \pi_n^{\mathbb{Q}}}{\prod_{n=0}^{N-1} \pi_n^{\mathbb{P}}} \right) \prod_{n=0}^{N-1} \pi_n^{\mathbb{Q}} = \mathbb{E}^{\mathbb{Q}} \left[\ln \left(\frac{\prod_{n=0}^{N-1} \pi_n^{\mathbb{Q}}}{\prod_{n=0}^{N-1} \pi_n^{\mathbb{P}}} \right) \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\ln \left(\prod_{n=0}^{N-1} \frac{\pi_n^{\mathbb{Q}}}{\pi_n^{\mathbb{P}}} \right) \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\sum_{n=0}^{N-1} \ln \left(\frac{\pi_n^{\mathbb{Q}}}{\pi_n^{\mathbb{P}}} \right) \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\sum_{n=0}^{N-1} \mathbb{E}^{\mathbb{Q}} \left[\ln \left(\frac{\pi_n^{\mathbb{Q}}}{\pi_n^{\mathbb{P}}} \right) \middle| S_n \right] \right] \quad (3.5)
 \end{aligned}$$

En particular consideramos aproximaciones inducidas por árboles trinomiales que permiten calcular la entropía relativa de forma explícita.

Sea $S_{n+1} = S_n H_{n+1}$, $n = 0, 1, \dots$, con $\Delta t = \frac{T}{N}$ y

$$H_{n+1} = \begin{cases} e^{\bar{\sigma}\sqrt{\Delta t}} & \text{con probabilidad } P_U \\ 1 & \text{con probabilidad } P_M \\ e^{-\bar{\sigma}\sqrt{\Delta t}} & \text{con probabilidad } P_D. \end{cases}$$

Observe que el logaritmo de S_n es una marcha aleatoria en el latice $\{\nu\bar{\sigma}\sqrt{\Delta t}, \nu \in \mathbb{Z}\}$.

Las probabilidades P_U , P_M y P_D vienen dadas por

$$\begin{aligned}
 P_U &= \frac{p}{2} \left(1 - \frac{\bar{\sigma}\sqrt{\Delta t}}{2} \right) + \frac{\mu\sqrt{\Delta t}}{2\bar{\sigma}} \\
 P_M &= 1 - p \\
 P_D &= \frac{p}{2} \left(1 + \frac{\bar{\sigma}\sqrt{\Delta t}}{2} \right) - \frac{\mu\sqrt{\Delta t}}{2\bar{\sigma}}
 \end{aligned}$$

siendo μ , $\bar{\sigma}$ y p parámetros que ajustan los 2 primeros momentos de los logaritmos de los procesos en (3.1).

De forma mas específica, denotando por μ_0 y p_0 los parámetros correspondientes a \mathbb{P} y por μ y p los correspondientes a \mathbb{Q} , son escogidos de modo tal que con las probabilidades P_U , P_M y P_D , y la media y la varianza de $\ln(S_n)$ vengán dadas por³ $\mu - \frac{p\bar{\sigma}}{2}$ y $p\bar{\sigma}^2$ respectivamente y con respecto a la medida de probabilidad correspondiente.

²Se puede considerar sólo el último punto S_n ya que este es un proceso de Markov.

³Nota: la suma de las esperanzas condicionales diverge cuando $\Delta t \rightarrow 0$

Seguindo la referencia [1], para Δt suficientemente pequeño tenemos que

$$\mathbb{E}^{\mathbb{Q}} \left[\ln \left(\frac{\pi_n^{\mathbb{Q}}}{\pi_n^{\mathbb{P}}} \right) \middle| S_n \right] = p \ln \left(\frac{p}{p_0} \right) + (1-p) \ln \left(\frac{1-p}{1-p_0} \right) + O(\Delta t) \quad (3.6)$$

Sustituyendo p y p_0 por $\sigma^2(t_n)/\bar{\sigma}^2$ y $\sigma_0^2/\bar{\sigma}^2$ respectivamente en la expresión (3.6) obtenemos

$$\mathbb{E}^{\mathbb{Q}} \left[\ln \left(\frac{\pi_n^{\mathbb{Q}}}{\pi_n^{\mathbb{P}}} \right) \middle| S_n \right] = \eta(\sigma^2(t_n)) + O(\Delta t), \quad (3.7)$$

con

$$\eta(\sigma^2(t_n)) := \frac{\sigma_n^2}{\bar{\sigma}^2} \ln \left(\frac{\sigma_n^2}{\sigma_0^2} \right) + \left(1 - \frac{\sigma_n^2}{\bar{\sigma}^2} \right) \ln \left(\frac{\bar{\sigma}^2 - \sigma_n^2}{\bar{\sigma}^2 - \sigma_0^2} \right) \quad (3.8)$$

Donde $\sigma_n^2 = \sigma^2(t_n)$, observe además que η se anula cuando $\sigma(t_n) = \sigma_0$. Sustituyendo (3.7) en (3.5) conseguimos

$$\begin{aligned} \mathbb{H}(\mathbb{Q}; \mathbb{P}) &= \mathbb{E}^{\mathbb{Q}} \left[\sum_{n=0}^{N-1} (\eta(\sigma^2(t_n)) + O(\Delta t)) \right] \\ &= \frac{1}{\Delta t} \mathbb{E} \left[\sum_{n=0}^{N-1} \eta(\sigma^2(t_n)) \Delta t \right] + O(1) \\ &\approx \frac{N}{T} \mathbb{E}^{\mathbb{Q}} \left[\int_0^T \eta(\sigma^2(t)) dt \right] + O(1). \end{aligned} \quad (3.9)$$

En esta última expresión podemos ver que para $\Delta t \ll 1$ suficientemente pequeño, la cantidad relevante en términos del contenido de información de la medida de probabilidad \mathbb{Q} con respecto a la medida de probabilidad \mathbb{P} es

$$\frac{1}{T} \mathbb{E}^{\mathbb{Q}} \left[\int_0^T \eta(\sigma(t))^2 dt \right], \quad (3.10)$$

(3.10) representa el límite de la entropía relativa por unidad de paso en el tiempo Δt (o tasa de entropía relativa) de la medida de probabilidad \mathbb{Q} con respecto a la medida \mathbb{P} . Vamos a modificar el problema de optimización (3.4) minimizando en vez esta cantidad sujeta a las restricciones (3.3) más adelante.

La noción de entropía relativa por unidad de tiempo no es una propiedad de los procesos de Itô (3.1), en el sentido de que la función $\eta(\cdot)$ no es única ya que depende de la discretización usada para aproximar los procesos. Es por lo tanto conveniente hacer una construcción en la que $\eta(\sigma^2)$ pertenezca a un grupo de funciones que que incluya a (3.8), así como a otras que pueden ser adecuadas.

Definición 3.3. Una función $\eta(\sigma^2)$ es una función de pseudo-entropía (**PE**) con mínimo en σ_0^2 si es suave, real valuada con $0 < \sigma^2 < +\infty$ y tal que

- i $0 \leq \eta(\sigma^2) < +\infty$
- ii $\eta(\sigma^2)$ es estrictamente convexa.
- iii $\eta(\sigma^2)$ tiene su valor mínimo, 0, en $\sigma^2 = \sigma_0^2$.

La función η definida anteriormente es una función de \mathbf{PE} , pero trabajaremos con una \mathbf{PE} más simple, con volatilidad a priori σ_0^2

$$\eta(\sigma^2) = \frac{1}{2} (\sigma_0^2 - \sigma^2)^2, \quad \sigma^2 \geq 0$$

La aproximación en (3.9) sugiere entonces que consideremos el siguiente problema de optimización

$$\inf_{\lambda} \sup_{\mathbb{Q}} \left\{ -\mathbb{E}^{\mathbb{Q}} \left[\int_0^T \eta(\sigma(t)) dt \right] + \sum_{i=1}^M \lambda_i \left(e^{-rT_i} \int G_i(s) dF_{T_i}^{\mathbb{Q}}(s) - C_i^{\text{Obs}} \right) \right\} \quad (3.11)$$

3.3. El Modelo de Heston y la Noción de Entropía Relativa Juntos.

En esta sección vamos a reescribir el problema de optimización (3.11) en el contexto del modelo de Heston de volatilidad estocástica

$$\begin{aligned} dS_t &= (\bar{r} - r_f) S_t dt + \sqrt{v_t} S_t [\sqrt{1 - \rho^2} dW_1 + \rho dW_2] \\ dv_t &= \kappa(\theta - v_t) dt + \sigma \sqrt{v_t} dW_2, \end{aligned} \quad (3.12)$$

donde ahora suponemos que $W_{1,t}$ y $W_{2,t}$ son movimientos Brownianos independientes.

La idea consiste en parametrizar el conjunto de medidas sobre las que se realiza la optimización en el problema (3.11) de acuerdo con el vector $\Theta = (\kappa, \theta, \sigma, \rho)$ de parámetros constantes del modelo de Heston. Es decir, consideramos una correspondencia tal que $\mathbb{P} = \mathbb{P}_{\Theta_0}$ es la medida de probabilidad del proceso (S_t, v_t) cuando el vector de parámetros en (3.12) es $\Theta_0 = (\kappa_0, \theta_0, \sigma_0, \rho_0)$ y $\mathbb{Q} = \mathbb{Q}_{\Theta}$ es la medida de probabilidad asociada a (S_t, v_t) cuando el vector de parámetros en (3.12) es $\Theta = (\kappa, \theta, \sigma, \rho)$. En términos de esta parametrización consideramos el problema de optimización

$$\inf_{\lambda} \sup_{\Theta} \left\{ -\mathbb{E} \left[\int_0^T (\bar{v} - v_s^{\Theta})^2 ds \right] + \sum_{i=1}^M \lambda_i \left(e^{-rT_i} \int G_i(s) dF_{T_i}^{\Theta}(s) - C_i^{\text{Obs}} \right) \right\} \quad (3.13)$$

teniendo en cuenta además que

$$e^{-rT} \int_{-\infty}^{\infty} G(s) dF_T^{\Theta}(s) = e^{-rT} \mathbb{E}[(S_T^{\Theta} - K)^+],$$

podemos reescribir (3.13) de la forma

$$\inf_{\lambda} \sup_{\Theta} \left\{ \mathbb{E} \left[- \int_0^T (\bar{v} - v_s^\Theta)^2 ds + \sum_{i=1}^M \lambda_i (e^{-rT_i} (S_{T_i}^\Theta - K_i)^+ - C_i^{\text{Obs}}) \right] \right\}. \quad (3.14)$$

Como veremos más adelante (3.14) es apropiada para aproximaciones numéricas. Antes de seguir adelante vamos a hacer varias observaciones referentes a esta última expresión:

- Primero obsérvese que desde la formula (3.13) ya dejamos de utilizar $\mathbb{E}^{\mathbb{Q}}[\cdot]$ para hacer referencia bajo cuál medida de probabilidad se está calculando la esperanza. En (3.14) el superíndice Θ en los proceso S_t y v_t indica que el valor esperado es calculado con respecto a la probabilidad \mathbb{Q}_Θ determinada cuando $\Theta = (\kappa, \theta, \sigma, \rho)$ es el vector de parámetros que rige la dinámica en (3.12).
- La constante \bar{v} que aparece en el término integral, recoge la información de la distribución a priori \mathbb{P}_{Θ_0} en forma tal que \bar{v} puede tomarse como una tasa de volatilidad promedio sobre el intervalo $[0, T]$ cuando el vector de parámetros $\Theta_0 = (\kappa_0, \theta_0, \sigma_0, \rho_0)$ rige la dinámica en (3.12).⁴

Finalmente para trabajar numéricamente la expresión en (3.14), vamos a definir la función V como sigue

$$V(S_{t_0}, v_{t_0}, t_0; \Theta, \lambda) = \mathbb{E} \left[- \int_0^T (\bar{v} - v_s^\Theta)^2 ds + \sum_{i=1}^M \lambda_i e^{-rT_i} (S_{T_i}^\Theta - K_i)^+ \middle| S_{t_0}, v_{t_0} \right] \quad (3.15)$$

donde como hemos venido asumiendo $\Theta = (\kappa, \theta, \sigma, \rho)$, $\lambda = (\lambda_1, \dots, \lambda_M)$, $r = \bar{r} - r_f$ y los $\{T_i\}$ son ahora tiempos al vencimiento vistos desde $t_0 = 0$. Consideramos entonces el problema

$$\begin{aligned} \inf_{\lambda} \sup_{\Theta} V(S_{t_0}, v_{t_0}, t_0; \Theta, \lambda) - \sum_{i=1}^M \lambda_i C_i^{\text{Obs}} = \\ \inf_{\lambda} \sup_{\Theta} V(S_{t_0}, v_{t_0}, t_0; \Theta, \lambda) - \lambda \cdot (C_1^{\text{Obs}}, \dots, C_M^{\text{Obs}}) \end{aligned} \quad (3.16)$$

⁴O sea $\bar{v} = \frac{1}{T} \mathbb{E}^{\mathbb{P}_{\Theta_0}} \left[\int_0^T v_s ds \right]$.

CAPÍTULO 4

Enfoque Numérico del Problema de Optimización

Primero veamos que la función $V(S_t, v_t, t)$ satisface cierto tipo de recurrencia en el tiempo como consecuencia de la propiedad de Markov de los procesos S_t y v_t , dicha relación de recurrencia permitirá diseñar aproximaciones numéricas para la función V .

Para comenzar consideramos una partición regular $0 = t_0 < t_1 < \dots < t_N = T$ con paso en el tiempo $\Delta t = \frac{T}{N}$ del intervalo $[0, T]$, donde $T = \max_i \{T_i\}_{i=1}^M$ es el mayor tiempo al vencimiento considerado, también vamos a suponer que todos los vencimientos considerados $\{T_i\}$ están incluidos en los puntos de la partición $\{t_k\}_{k=0}^N$.

Retomamos la función V del final del capítulo anterior, pero omitiremos los parámetros Θ y λ para relajar un poco la notación. Ahora, $V(S_t, v_t, t)$ para $0 \leq t \leq T$ será

$$V(S_t, v_t, t) = \mathbb{E} \left[- \int_t^T (\bar{v} - v_s^\Theta)^2 ds + \sum_{i:t \leq T_i} \lambda_i e^{-rT_i} (S_{T_i}^\Theta - K_i)^+ \middle| S_t, v_t \right]. \quad (4.1)$$

Supongamos que los 2 últimos vencimientos son T_{M-1} y $T_M = T$, observando además que

$$V(S_T, v_T, T) = \lambda_M e^{-rT} (S_T - K_M)^+, \quad (4.2)$$

condicionamos la expresión interna de la esperanza en (4.1) con respecto a $(S_{T-\Delta t}, v_{T-\Delta t})$ para Δt pequeño tenemos

$$\begin{aligned} V(S_{T-\Delta t}, v_{T-\Delta t}, T - \Delta t) &= \mathbb{E}_{T-\Delta t} \left[- \int_{T-\Delta t}^T (\bar{v} - v_s)^2 ds + \lambda_M e^{-rT} (S_T - K_M)^+ \right] \\ &= \mathbb{E}_{T-\Delta t} \left[- \int_{T-\Delta t}^T (\bar{v} - v_s)^2 ds + V(S_T, v_T, T) \right], \end{aligned} \quad (4.3)$$

donde $\mathbb{E}_{T-\Delta t} := \mathbb{E}[\cdot | S_{T-\Delta t}, v_{T-\Delta t}]$.

Del mismo modo si para Δt pequeño condicionamos la expresión interna en (4.1) esta vez con respecto a $(S_{T_{M-1}^+}, v_{T_{M-1}^+})$, podemos ver que¹

$$\begin{aligned}
V(T_{M-1} - \Delta t) &= \mathbb{E}_{T_{M-1} - \Delta t} \left[\mathbb{E} \left[- \int_{T_{M-1} - \Delta t}^T (\bar{v} - v_s)^2 ds + \lambda_{M-1} e^{-rT_{M-1}} (S_{T_{M-1}} - K_{M-1})^+ \right. \right. \\
&\quad \left. \left. + \lambda_M e^{-rT} (S_{T_M} - K_M)^+ \middle| \mathcal{F}_{T_{M-1}^+} \right] \right] \\
&= \mathbb{E}_{T_{M-1} - \Delta t} \left[- \int_{T_{M-1} - \Delta t}^{T_{M-1}} (\bar{v} - v_s)^2 ds + \lambda_{M-1} e^{-rT_{M-1}} (S_{T_{M-1}} - K_{M-1})^+ \right. \\
&\quad \left. + \mathbb{E} \left[- \int_{T_{M-1}}^T (\bar{v} - v_s)^2 ds + \lambda_M e^{-rT} (S_T - K_M)^+ \middle| \mathcal{F}_{T_{M-1}^+} \right] \right] \\
&= \mathbb{E}_{T_{M-1} - \Delta t} \left[- \int_{T_{M-1} - \Delta t}^{T_{M-1}} (\bar{v} - v_s)^2 ds + \lambda_{M-1} e^{-rT_{M-1}} (S_{T_{M-1}} - K_{M-1})^+ \right. \\
&\quad \left. + V(T_{M-1}^+) \right].
\end{aligned} \tag{4.4}$$

Donde hemos abusado levemente de la notación escribiendo $V(t)$ para $V(S_t, v_t, t)$. En conclusión, la relación de recurrencia satisfecha por la función V a la que nos hemos referido con anterioridad y que está implícita en las ecuaciones (4.3) y (4.4) puede escribirse como

$$V(t_{k-1}) = \mathbb{E} \left[- \int_{t_{k-1}}^{t_k} (\bar{v} - v_s) ds + \delta_{\{t_k=T_j\}} \lambda_j e^{-rT_j} (S_{T_j} - K_j)^+ + V(t_k^+) \right], \tag{4.5}$$

donde evaluamos V en los tiempos t_n de la partición de $[0, T]$ con Δt suficientemente pequeño, la delta de Kroneker $\delta_{\{t_k=T_j\}}$ asegura que el término $e^{-rT_j} (S_{T_j} - K_j)^+$ aparece en (4.5) si y solo si t_k coincide con alguno de los vencimientos T_j .

La idea ahora es aproximar (4.5) por medio de una ecuación en diferencias que será conseguida mediante una aproximación adecuada del modelo de Heston por medio de un árbol recombinante.

4.1. Aproximación del modelo de Heston

En esta sección vamos a reproducir las ideas propuestas en [3] de utilizar un árbol recombinante para aproximar las difusiones de Itô que caracterizan el modelo de Heston, de forma más precisa retomemos el modelo de Heston

¹ t^+ significa límite por la derecha en t y $\mathcal{F}_{T_{M-1}^+}$ se trata de la σ -álgebra $\mathcal{F}_{T_{M-1}^+} = \bigcap_{\{t: T_{M-1} < t\}} \sigma(\{S_t, v_t\})$.

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{v_t} S_t dW_1 \\ S_{t_0} &= S_0. \end{aligned}$$

$$\begin{aligned} dv_t &= \kappa(\theta - v_t) dt + \sigma \sqrt{v_t} dW_2 \\ v_{t_0} &= v_0, \end{aligned}$$

(4.6)

donde W_1 y W_2 son movimientos brownianos con correlación instantánea ρ , i.e. $\mathbb{E}[dW_1 dW_2] = \rho dt$. La idea es trabajar con dos procesos que tengan movimientos brownianos independientes, consideramos entonces la siguiente transformación sobre los proces S_t y v_t

$$x_t := \ln(S_t), \quad y_t := \frac{v_t}{\sigma} - \rho x_t \quad (4.7)$$

y usando la fórmula de Itô se obtiene

$$\begin{aligned} dx_t &= \mu_x(x_t, y_t) dt + \sqrt{\sigma} \phi(x_t, y_t) dW_t \\ x_0 &= \ln(S_0). \end{aligned} \quad (4.8)$$

$$\begin{aligned} dy_t &= \mu_y(x_t, y_t) dt + \sqrt{\sigma(1-\rho^2)} \phi(x_t, y_t) d\bar{W}_t \\ y_0 &= \frac{v_0}{\sigma} - \rho x_0, \end{aligned} \quad (4.9)$$

(4.10)

donde²

$$\begin{aligned} \mu_x(x, y) &:= r - \frac{\sigma}{2}(y + \rho x) \\ \mu_y(x, y) &:= \frac{\kappa\theta}{\sigma} - r\rho + \frac{1}{2}(\rho\sigma - 2\kappa)(y + \rho x) \\ \phi(x, y) &:= \sqrt{(y + \rho x)^+} \\ dW &:= dW_1, \quad d\bar{W}_t := \frac{dW_2 - \rho dW_1}{\sqrt{1-\rho^2}} \end{aligned}$$

Ahora vamos a considerar marchas aleatorias discretas $\{x_k^{(N)}, y_k^{(N)}\}_{k=0}^N$, que discretizan tanto el tiempo como los valores que toma el subyacente.³

$$\begin{aligned} x_k^{(N)} &:= x_0 + \sqrt{h\sigma} \sum_{i=1}^k \xi_i^x \quad \text{con} \quad x_0 = \ln(S_0) \\ y_k^{(N)} &:= y_0 + \sqrt{h\sigma(1-\rho^2)} \sum_{i=1}^k \xi_i^y \quad \text{con} \quad y_0 = \frac{v_0}{\sigma} - \rho x_0 \end{aligned}$$

²La función $\eta(x, y)$ no es la misma de la que se habló en la sección anterior en dónde se le tomaba como función de **PE**

³ N denota el número de elementos de la partición.

donde $h = \Delta t = \frac{T}{N}$. Los procesos $\{x_k^{(N)}, y_k^{(N)}\}$ aproximan (4.8) y (4.9) respectivamente, (ξ^x, ξ^y) son variables aleatorias Bernoulli que toman valores en $\{1, -1\}$ y $\xi_0 := 0$.

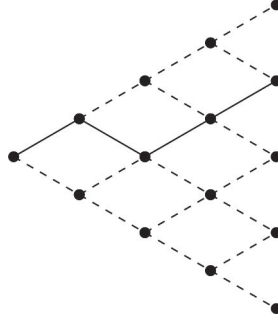


FIGURA 4.1. Marchas discretas.

Una vez más las marchas aleatorias $\{x_k^{(N)}, y_k^{(N)}\}$ son modificadas de manera tal que la tupla $(x_k^{(N)}, y_k^{(N)}, \xi_k^x, \xi_k^y)$ constituya una cadena de Markov⁴ de la siguiente forma

$$\begin{aligned}\bar{x}_k^{(N)} &:= x_k^{(N)} + \sqrt{h\sigma}\alpha_k\xi_k^x \\ \bar{y}_k^{(N)} &:= y_k^{(N)} + \sqrt{h\sigma(1-\rho^2)}\beta_k\xi_k^y, \text{ con } k = 1, \dots, N\end{aligned}$$

En donde los procesos $\{\alpha_k, \beta_k\}_{k=1}^N$ son construidos de forma tal que los incrementos de $\{\bar{x}_k^{(N)}, \bar{y}_k^{(N)}\}$ ajusten los 2 primeros momentos de los incrementos de (x_t, y_t) . Y así, el proceso $\{\bar{x}_k, \bar{y}_k\}$ deberá cumplir las siguientes propiedades:

$$\mathbb{E}_{k-1} [\bar{x}_k - \bar{x}_{k-1}] = \mu_x(x_{k-1}, y_{k-1})h + o(h) \quad (4.11)$$

$$\mathbb{E}_{k-1} [\bar{y}_k - \bar{y}_{k-1}] = \mu_y(x_{k-1}, y_{k-1})h + o(h) \quad (4.12)$$

$$\mathbb{E}_{k-1} [(\bar{x}_k - \bar{x}_{k-1})^2] = \sigma\phi^2(x_{k-1}, y_{k-1})h + o(h) \quad (4.13)$$

$$\mathbb{E}_{k-1} [(\bar{y}_k - \bar{y}_{k-1})^2] = \sigma(1-\rho^2)\phi^2(x_{k-1}, y_{k-1})h + o(h) \quad (4.14)$$

$$(4.15)$$

donde

$$\mathbb{E}_{k-1} [\cdot] := \mathbb{E} [\cdot | \bar{x}_{k-1}, \bar{y}_{k-1}]$$

las probabilidades

$$p_k := \mathbb{P}_{k-1}(\xi_k^x = 1) \text{ y } q_k := \mathbb{P}_{k-1}(\xi_k^y = 1)$$

son calculadas escribiendo de forma explicita la siguiente relación

$$\mathbb{E}_{k-1} = \left[\exp(-rh) \exp(\bar{x}_k^{(N)}) - \exp(\bar{x}_{k-1}^{(N)}) \right] = 0 \quad (4.16)$$

⁴La convergencia de x y y es equivalente a la de \bar{x} y \bar{y} .

(4.16) expresa que el precio descontado es una martingala (condición de no arbitraje) que al parecer mejora la convergencia de la discretización al ser usada en vez de (4.11).

Los detalles pueden ser consultados en [3], donde se obtienen las siguientes expresiones para las probabilidades p_k q_k que rigen la dinámica de la cadena de Markov $\{\bar{x}_k^{(N)}, \bar{y}_k^{(N)}\}$

$$\begin{aligned} p_k &= \left(\min 1, \frac{\exp(rh + \sqrt{h\sigma}\alpha_{k-1}\xi_{k-1}^x) - \exp(-\sqrt{h\sigma}(1 + \alpha_k))}{\exp(\sqrt{h\sigma}(1 + \alpha_k)) - \exp(-\sqrt{h\sigma}(1 + \alpha_k))} \right)^+ \\ q_k &= \left(\min \left\{ 1, \frac{1}{2} + \frac{\alpha_{k-1}}{2(1 + \alpha_k)}\xi_{k-1}^y + \frac{\sqrt{h}\mu_y(x_{k-1}, y_{k-1})}{2\sqrt{\sigma(1 - \rho^2)}(1 + \alpha_k)} \right\} \right)^+ \\ \alpha_k &= \beta_k = \frac{\phi^2(x_{k-1}, y_{k-1}) - 1}{2}. \end{aligned}$$

4.2. Implementación del Árbol Recombinante

Construiremos un enmallado $(\mathcal{X}_k, \mathcal{Y}_k)$ que definiremos enseguida para calcular V de forma descendente en el tiempo siguiendo los procesos x_k y y_k , que reescribiremos como

$$\begin{aligned} \mathcal{X}_k(l) &= x_0 + (2l - k)\sqrt{\sigma h} \quad \text{con } 0 \leq l \leq k \\ \mathcal{Y}_k(m) &= y_0 + (2m - k)\sqrt{\sigma h(1 - \rho^2)} \quad \text{con } 0 \leq m \leq k. \end{aligned}$$

De igual forma cambiamos la manera de escribir p_k ⁵ y q_k en términos de l y m

$$\begin{aligned} p_k(l, m, \xi^x, \xi^y) &= \left(\min \left\{ 1, \frac{\exp(rh + \sqrt{h\sigma}(\Psi_{k-1})\xi^x) - \exp(-\sqrt{h\sigma}\Psi_k)}{\exp(\sqrt{h\sigma}\Psi_k) - \exp(-\sqrt{h\sigma}\Psi_k)} \right\} \right)^+ \\ q_k(l, m, \xi^x, \xi^y) &= \left(\min \left\{ 1, \frac{1}{2} + \frac{\alpha_{k-1}(l - \xi^x, m - \xi^y)}{2\Psi_k}\xi^y + \frac{\sqrt{h}\mu_y(x_{k-1}, y_{k-1})}{2\sqrt{\sigma(1 - \rho^2)}\Psi_k} \right\} \right)^+ \end{aligned}$$

donde

$$\begin{aligned} \alpha_k(l, m) &= \frac{\phi^2(x_{k-1}(l), y_{k-1}(m)) - 1}{2} \\ \Psi_k &= 1 + \alpha_k(l, m) \\ \mu_y(x, y) &= \frac{\kappa\theta}{\sigma} - r\rho + \frac{1}{2}(\rho\sigma - 2\kappa)(y + \rho x) \end{aligned}$$

Consideramos una sucesión de funciones

$$V_n^{(N)}(l, m, \xi^x, \xi^y) := V^{(N)}(\mathcal{X}_n(l), \mathcal{Y}_n(m), \xi^x, \xi^y),$$

definidas sobre la malla $(\mathcal{X}_n, \mathcal{Y}_n)$ y determinadas por la relación de recurrencia

⁵ Agregamos el cálculo de un mínimo y de una parte positiva a p_k para garantizar que sea una probabilidad.

$$V_{k-1}(l, m, \xi^x, \xi^y) = \mathbb{E}_{k-1} \left[-\Lambda(\bar{x}_k, \bar{y}_k) \Delta t + \delta_{\{t_k=T_j\}} \lambda_j e^{-rT_j} (e^{\bar{x}_k} - K_j)^+ + V_k(\bar{x}_k, \bar{y}_k, \xi_k^x, \xi_k^y) \right], \quad (4.17)$$

donde $\Lambda(x, y) = (\bar{v} - \sigma(\rho x + y))^2$ aparece en el término que se encarga de aproximar la integral $\int_{t_{k-1}}^{t_k} (\bar{v} - v_s) ds$. Cuando t_k no coincide con ninguno de los vencimientos T_j en (4.17) simplemente tenemos

$$V_{k-1}(l, m, \xi^x, \xi^y) = \mathbb{E}_{k-1} \left[-\Lambda(\bar{x}_k, \bar{y}_k) \Delta t + V_k(\bar{x}_k, \bar{y}_k, \xi_k^x, \xi_k^y) \right]. \quad (4.18)$$

Calculando el valor esperado (4.17) en términos de las probabilidades de transición p_k y q_k tenemos que

$$\begin{aligned} V_{k-1}(l, m, \xi^x, \xi^y) = & \left(-\Lambda_k(l+1, m+1) \Delta t + \delta_{\{t_k=T_j\}} \lambda_j e^{-rT_j} (e^{\mathcal{X}_k(l+1)} - K_j)^+ + V_k(l+1, m+1, 1, 1) \right) p_k q_k \\ & + \left(-\Lambda_k(l+1, m) \Delta t + \delta_{\{t_k=T_j\}} \lambda_j e^{-rT_j} (e^{\mathcal{X}_k(l+1)} - K_j)^+ + V_k(l+1, m, 1, -1) \right) p_k (1 - q_k) \\ & + \left(-\Lambda_k(l, m+1) \Delta t + \delta_{\{t_k=T_j\}} \lambda_j e^{-rT_j} (e^{\mathcal{X}_k(l)} - K_j)^+ + V_k(l, m+1, -1, 1) \right) (1 - p_k) q_k \\ & + \left(-\Lambda_k(l, m) \Delta t + \delta_{\{t_k=T_j\}} \lambda_j e^{-rT_j} (e^{\mathcal{X}_k(l)} - K_j)^+ + V_k(l, m, -1, -1) \right) (1 - p_k) (1 - q_k). \end{aligned} \quad (4.19)$$

En (4.19) $\Lambda_k(l, m) := \Lambda(\mathcal{X}_k(l), \mathcal{Y}_k(m))$, también anotamos que (4.19) se resuelve hacia atrás en el tiempo notando que

$$V_N(l, m, \xi^x, \xi^y) = e^{-rT} (e^{\mathcal{X}_N(l)} - K_M)^+ = e^{-rT} (S_T - K_M)^+. \quad (4.20)$$

Debemos tener que

$$V_0^{(N)}(x_0, y_0) \rightarrow V(S_{t_0}, v_{t_0}, t_0), \quad \text{cuando } N \rightarrow \infty. \quad (4.21)$$

De la misma forma podemos aproximar los precios de opciones en el instante t_0

$$C(S_{t_0}, v_{t_0}, t_0; T_j, K_j) = e^{-rT_j} \mathbb{E}[(S_{T_j} - K_j)^+], \quad (4.22)$$

por medio de funciones $C_n(l, m, \xi^x, \xi^y)$ definidas en la malla $(\mathcal{X}_n, \mathcal{Y}_n)$ que cumplen la relación de recurrencia

$$\begin{aligned}
C_{k-1}(l, m, \xi^x, \xi^y) &= \mathbb{E}[C_k(\bar{x}_k, \bar{y}_k, \xi_k^x, \xi_k^y)], \quad k = N-1, N-2 \dots 0, \\
C_N(l, m, \xi^x, \xi^y) &= e^{-rT}(e^{\mathcal{X}_N(l)} - K_j)^+
\end{aligned} \tag{4.23}$$

Debemos tener que

$$C_0^N(l, m) \rightarrow C(S_{t_0}, v_{t_0}, t_0), \quad \text{cuando } N \rightarrow \infty \tag{4.24}$$

Para la minimización de la función V^6 necesitaremos de un optimizador global, el cuál pueda entregarnos el valor del mínimo para todo el conjunto en en cuál buscaremos nuestro vector de parámetros. Minimizaremos con un método basado en un fenómeno físico de cristalización que sufren los metales al ser enfriados desde un estado de fundición, el cual es llamado ASA por sus siglas en inglés (Adaptive Simulated Annealing). Para mayor profundidad se puede consultar la referencia [7].

4.3. Optimización de λ y Θ

A continuación describiremos, en términos generales, cómo pretendemos resolver el problema de optimización

$$\begin{aligned}
\inf_{\lambda} \sup_{\Theta} V(S_0, V_0; \Theta, \lambda) - \lambda \cdot C^{\text{Obs}} &= \\
&= \inf_{\lambda} \sup_{\Theta} \mathbb{E} \left[- \int_0^T (\bar{v} - v_t^{\Theta})^2 dt \right. \\
&\quad \left. + \sum_i \lambda_i \left(e^{-rT_i} (S_{T_i} - K_i)^+ - C_i^{\text{Obs}} \right) \middle| S_0, v_0 \right] \\
&= \inf_{\lambda} \sup_{\Theta} \mathbb{E} \left[- \int_0^T (\bar{v} - v_t^{\Theta})^2 dt \middle| S_0, v_0 \right] \\
&\quad + \sum_i \lambda_i \left(C_H(S_0, v_0; T_i, K_i, \Theta) - C_i^{\text{Obs}} \right)
\end{aligned} \tag{4.25}$$

nuevamente abusamos de manera leve en la notación en el sentido de que en los cálculos que son presentados a continuación utilizamos $V(S_{t_0}, v_{t_0}, t_0)$ pero realmente, se trata de la aproximación $V_0(x_0, y_0; \Theta, \lambda)$, de igual forma utilizamos en las fórmulas $C_H(S_0, v_0; T_i, K_i, \Theta)$ cuando se trata de las aproximaciones $C_0(x_0, y_0; \Theta)$.

En (4.25) debemos recordar que la función $C_H(S_0, v_0, T, K; \Theta)$ denota el precio de una opción europea con precio de ejercicio K en el vencimiento T cuando los parámetros poseen los valores dados por Θ . La optimización interior, o sea el sup en Θ es llevada a cabo por medio de ASA, haciendo uso del árbol recombinante para el cálculo de la función $V(S_0, v_0; \Theta, \lambda)$ para valores fijos de Θ y λ .

⁶Como fue escrita finalmente en este capítulo.

Denotando por

$$W(S_0, v_0; \Theta, \boldsymbol{\lambda}) = V(S_0, V_0; \Theta, \boldsymbol{\lambda}) - \boldsymbol{\lambda} \cdot \mathbf{C}^{\text{Obs}},$$

podemos observar que el gradiente en $\boldsymbol{\lambda}$ de la función $W(\cdot; \boldsymbol{\lambda})$ es

$$\nabla_{\boldsymbol{\lambda}} W(S_0, v_0; \Theta, \boldsymbol{\lambda}) = \mathbf{r} = (r_1, r_2, \dots, r_M)$$

donde \mathbf{r} es el vector de residuos con componentes r_i

$$r_i = C_H(S_0, v_0, T_i, K_i; \theta) - C_i^{\text{Obs}}$$

el cálculo de estos residuos es un subproducto del cálculo de la función V por medio del árbol recombinante.

La optimización exterior ínf en $\boldsymbol{\lambda}$ será tratada por medio del método del gradiente

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + h_k \frac{\nabla_{\boldsymbol{\lambda}} W(\boldsymbol{\lambda}_k)}{\|\nabla_{\boldsymbol{\lambda}} W(\boldsymbol{\lambda}_k)\|} \quad k = 0, 1, 2 \dots \quad (4.26)$$

donde la secuencia de pasos (h_k) se escoge de forma adecuada tal que $\lim h_k = 0$.

Resumen del Algoritmo

Paso 1 . Proponer valores iniciales $\Theta_0, \boldsymbol{\lambda}_0$ para Θ y $\boldsymbol{\lambda}$.

Paso 2 . El algoritmo de temple simulado adaptativo (ASA) permite tomar Θ_k como

$$\Theta_k = \arg \max_{\Theta} W(\boldsymbol{\lambda}_k, \Theta).$$

simultáneamente se calculan los residuos $r_i(\Theta_k) = C_H(S_0, v_0, T_i, K_i; \Theta_k) - C_i^{\text{Obs}}$. Por medio del árbol binomial.

Paso 3 . Resolver para h_k el paso del método del gradiente

$$h_k = \arg \min_h W(\boldsymbol{\lambda}_k + h \frac{\mathbf{r}_k}{\|\mathbf{r}_k\|}, \Theta_k). \quad (4.27)$$

Paso 4 Actualizar $\boldsymbol{\lambda}$

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + h_k \frac{\mathbf{r}_k}{\|\mathbf{r}_k\|}. \quad (4.28)$$

El algoritmo termina si digamos

$$\frac{\|\boldsymbol{\lambda}_{k+1} - \boldsymbol{\lambda}_k\|}{\|\boldsymbol{\lambda}_k\|} < \epsilon_{\text{tol}}$$

en otro caso ($\frac{\|\boldsymbol{\lambda}_{k+1} - \boldsymbol{\lambda}_k\|}{\|\boldsymbol{\lambda}_k\|} > \epsilon_{\text{tol}}$) volver al Paso 2.

Cuando el último $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}_{m+1}$ calculado es tal que $\frac{\|\boldsymbol{\lambda}_{m+1} - \boldsymbol{\lambda}_m\|}{\|\boldsymbol{\lambda}_m\|} < \epsilon_{\text{tol}}$ debemos tener que el par $(\boldsymbol{\lambda}^*, \Theta^*) = (\boldsymbol{\lambda}_{m+1}, \Theta_m)$ debe ser una aproximación razonable del problema de optimización (4.25), i.e.

$$W(S_0, v_0; \boldsymbol{\lambda}^*, \Theta^*) \approx \inf_{\boldsymbol{\lambda}} \sup_{\Theta} W(S_0, v_0; \boldsymbol{\lambda}, \theta). \quad (4.29)$$

El vector de parámetros $\Theta^* = (\kappa^*, \theta^*, \sigma^*, \rho^*)$ es el candidato para ajustar el modelo. Es decir, la distribución del proceso $S_t^{\Theta^*}$ donde

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{v_t} S_t (\sqrt{1 - \rho^{*2}} dW_{1,t} + \rho^* dW_{2,t}) \\ dv_t &= \kappa^* (\theta^* - v_t) dt + \sigma^* \sqrt{v_t} dW_{2,t}. \end{aligned} \quad (4.30)$$

minimiza la entropía con respecto a la distribución a priori y satisface

$$C_H(S_0, v_0, T_i, K_i; \Theta^*) \approx C_i^{\text{Obs}}, \quad i = 1, \dots, M.$$

para todo el conjunto de precios de opciones en el mercado $\{C_i^{\text{Obs}}\}$.

Resultados

5.1. Árbol Recombinante

Con las ideas del árbol recombinante construimos funciones en MATLAB las que maximizamos respecto a las variables κ , θ , σ y ρ . El conjunto de λ 's que tenían que hacer parte del proceso de optimización fueron tomados constantes¹ pero se repitió la maximización para varios valores de estos. El conjunto de valores para estos, se tomó como un vector columna de 1's, multiplicados por una constante β que varió entre 0.00001 y 1². Los resultados se consignan en la siguiente tabla, en la cuál se discretizó en 20 subintervalos, donde la valor V0 representa $V_0(0, 0, 0, 0)$ y toc el tiempo de ejecución de la optimización

β	κ	θ	σ	ρ	V0	toc
1	8,0293	0,3511	0,8553	0,9836	61,9797	772,9374
0,1	2,6130	0,8934	0,9732	0,9881	5,0642	780,2973
0,01	9,0844	0,0350	0,0782	0,9771	-0,6246	771,7203
0,001	9,0844	0,0350	0,0782	0,9771	-1,1938	776,8516
0,0001	2,6130	0,8934	0,9732	0,9881	-1,2535	774,9725
0,00001	6,8124	0,9943	0,9838	0,9963	-1,2628	758,9884

TABLA 5.1. Resultados del proceso de calibración con λ variable y con 20 pasos temporales.

También se realizó un experimento adicional con 40 subintervalos de tiempo, por cuestiones de costo tiempo de ejecución se hizo sólo este

β	κ	θ	σ	ρ	V0	toc
1	8.3516	0.6565	0.9846	0.9596	61.9124	85645

TABLA 5.2. Resultados del proceso de calibración con 40 pasos temporales.

¹La inclusión de el conjunto de λ 's en el proceso de optimización es algo que podría realizarse en un trabajo posterior.

² $\lambda = \beta \cdot \text{ones}(m, 1)$, donde m es el número de datos.

Los datos que fueron usados para este conjunto de ensayos fueron tomados el 14 de julio del año 2014 de www.barchart.com y constan de un vencimiento de 53 días. Estos se consignan en la siguiente tabla en la que K simboliza los strikes y C los precios de las opciones observados

K	C	K	C	K	C	K	C	K	C
1.2	0.16220	1.265	0.09720	1.33	0.03380	1.395	0.00090	1.46	0.00001
1.205	0.15720	1.27	0.09220	1.335	0.02930	1.4	0.00060	1.465	0.00001
1.21	0.15220	1.275	0.08730	1.34	0.02500	1.405	0.00040	1.47	0.00001
1.215	0.14720	1.28	0.08230	1.345	0.02090	1.41	0.00030	1.475	0.00001
1.22	0.14220	1.285	0.07730	1.35	0.01710	1.415	0.00020	1.48	0.00001
1.225	0.13720	1.29	0.07230	1.355	0.01370	1.42	0.00015	1.485	0.00001
1.23	0.13220	1.295	0.06740	1.36	0.01060	1.425	0.00010	1.49	0.00001
1.235	0.12720	1.3	0.06240	1.365	0.00800	1.43	0.00010	1.495	0.00001
1.24	0.12220	1.305	0.05750	1.37	0.00590	1.435	0.00005	1.5	0.00001
1.245	0.11720	1.31	0.05260	1.375	0.00420	1.44	0.00005	1.505	0.00001
1.25	0.11220	1.315	0.04780	1.38	0.00290	1.445	0.00001	1.51	0.00001
1.255	0.10720	1.32	0.04310	1.385	0.00200	1.45	0.00001	1.515	0.00001
1.26	0.10220	1.325	0.03840	1.39	0.00130	1.455	0.00001	1.52	0.00001

TABLA 5.3. Precios de opciones y strikes que se usaron en la calibración.

5.2. Regularización Vía Tychonoff

Para diferentes valores de α se realizó el siguiente problema de optimización:

$$\min_{\Theta} \left\{ \|C(S_0, v_0, t_0; \Theta) - C^{mk}\| + \alpha \|\Theta - \Theta_0\| \right\} \quad (5.1)$$

Para realizar la calibración se usaron los siguientes datos: $s_0 = 1,3358$ es la tasa de cambio de febrero 15 de 2013, y $r = 0,02$ y $v_0 = 0,5$ y se tomaron procurando tomar valores razonables. Los parámetros *strike*, que simbolizaremos con K , y T son obtenidos de un conjunto de valores de opciones call para la fecha indicada anteriormente y de igual forma el vector C^{mk} que aparece en 5.1 se obtuvo de tal tabla anterior.

Observación 5.1. Como en fue realizado en para el árbol recombinante, para este proceso también se usó el mismo principio de optimización.

Los resultados fueron los siguientes donde **toc** será el tiempo, en segundos, que demoró el computador en obtener un resultado y **fstar** es el valor de $\|C(S_0, v_0, t_0; \Theta) - C^{mk}\| + \alpha \|\Theta - \Theta_0\|$ evaluada en el *kappa*, *theta*, *sigma* y *rho* obtenidos

α	fstar	κ	θ	σ	ρ	toc
0	0,2560	9,6600	0,0194	0,6046	-0,9622	2398,4758
0,001	0,3520	9,7061	0,0024	0,1482	-0,8071	2457,3529
0,01	0,5638	2,3091	0,0120	0,1672	-0,5812	2403,7553
0,1	0,6357	0,4895	0,0286	0,0580	-0,1580	2347,7202
1	0,6891	0,0181	0,2056	0,0652	0,0694	2315,5925
10	1,2045	0,1029	0,1632	0,1409	0,0086	2370,3207
100	3,4913	0,0434	0,1540	0,0541	0,0027	2362,9506
1000	158,4893	0,2416	0,2416	0,2416	0,2416	2337,5545

TABLA 5.4. Resultados del experimento.

Conclusiones

- El costo en tiempo de la calibración de parámetros puede llegar a ser muy alto si se pretende tener un grado alto de exactitud.
- Deben buscarse métodos que sean más rápidos, o bien, máquinas más potentes para continuar la investigación en esta área.
- El problema de calibración resultó ser un amplio tema de investigación, y preferimos dejar el problema de la cobertura para trabajos posteriores.
- Puede ser viable realizar procesos de calibración en periodos de tiempo razonable: una o dos semanas.
- La minimización de entropía relativa y la regularización de Tychonoff mostraron un gasto temporal similar en algunos experimentos.

Definición Intuitiva de la Entropía

La idea de entropía viene de conceptos físicos, la cual mide “el desorden” en un sistema, ahora bien, la forma de asignarle una cantidad a tal concepto recae en cómo se describe la situación, esto nos lleva a pensar una forma de cuantificar qué tanta información es requerida para tal fin. Prosigamos pues a definir la entropía como será usada en nuestro estudio. Supongamos que tenemos un evento, la manera de describirlo requeriría información, y sería necesaria más de esta mientras más extraño sea, por lo cual deberíamos encontrar una forma de medir la cantidad información y está basada en cómo se acostumbra guardarla. Un buen método para este fin es el de pensar las descripciones como cadenas de bits, 1's y 0's, así que un mensaje con mucha información sería más largo, por ende, al pensarlo como un número en sistema binario sería un número más grande. Y si al pensar que a cada mensaje se le asigna un número entero positivo la medida de la información estaría dada por:

$$\log_2(x), \text{ donde } x \in \mathbb{Z}$$

Denotemos por H a la entropía y para definirla en nuestro contexto, comencemos con una variable aleatoria X que tiene n resultados posibles x_1, x_2, \dots, x_n los cuales son igualmente probables, así la función de densidad de probabilidad de X es $P(X = x_i) = p = 1/n$ con $i = 1, \dots, n$ y diremos que la entropía de una variable aleatoria vendrá dada por la cantidad de información requerida para describirlo, i.e.

$$\begin{aligned} H(X) &= \log_2(n) \\ &= \log_2\left(\frac{1}{p}\right) \\ &= \log_2(1) - \log_2(p) \\ &= -\log_2(p) = \frac{n}{n}(-\log_2(p)) \\ &= \sum_{i=1}^n \frac{1}{n}(-\log_2(p)) \\ &= \sum_{i=1}^n P(X = x_i)(-\log_2(P(X = x_i))) \\ &= \mathbb{E}[-\log_2(P(X))] \end{aligned}$$

APÉNDICE B

Entropía Relativa en el caso de un solo vencimiento

Si vemos a todo lo que está bajo la acción del $\text{sup}_{\mathbb{Q}}$ en 3.4 como un funcional que toma a f , donde $d\mathbb{Q} = f dx$, y que calcula:

$$-\int f \ln\left(\frac{f}{f_0}\right) dx + \sum_i \lambda_i \left(\int G_i(x) f(x) dx - C_i \right) + \mu \left(\int f dx - 1 \right)$$

encontrar dicho sup se reduce a hallar cuando la derivada en alguna dirección δf se anule, la formulación para esto está dada así:

$$\begin{aligned} \int \delta f \ln\left(\frac{f}{f_0}\right) - \delta f + \sum_i \lambda_i \delta f G_i + \mu \delta f &= \\ \int \delta f \left(\ln\left(\frac{f}{f_0}\right) - 1 + \sum_i \lambda_i G_i + \mu \right) &= 0 \end{aligned}$$

donde la función δf se escoge de manera arbitraria y como queremos que la integral se anule, se tiene que la distribución f debe cumplir

$$-\ln\left(\frac{f}{f_0}\right) - 1 + \sum_i \lambda_i G_i + \mu = 0$$

Así, despejando a f ¹

$$\begin{aligned} -\ln\left(\frac{f_\lambda}{f_0}\right) &= 1 - \sum_i \lambda_i G_i - \mu \Leftrightarrow \\ f_\lambda &= f_0 \exp\left(\mu - 1 + \sum_i \lambda_i G_i\right) \Leftrightarrow \\ \int f_\lambda &= 1 \Leftrightarrow e^{\mu-1} \int f_0 e^{\lambda \cdot G} = 1 \end{aligned} \tag{B.1}$$

¹La notación f_λ enfatiza la dependencia de f con respecto al vector $\lambda = (\lambda_1, \dots, \lambda_N)$.

Ergo,

$$e^{\mu-1} = \frac{1}{\int f_0 e^{\lambda \cdot G}} \quad (\text{B.2})$$

y definiendo a la función $Z(\lambda)$,

$$\int f_0 e^{\lambda \cdot G} := Z(\lambda) \quad (\text{B.3})$$

Obtenemos una expresión para f_λ dependiendo únicamente del vector de λ 's.

$$f_\lambda = \frac{f_0 e^{\lambda \cdot G}}{Z(\lambda)}$$

Acabamos de replantear el problema de hallar una función óptima (f_λ) a encontrar un conjunto de parámetros que la definen, esto es mucho más simple pues traducimos una cuestión infinito dimensional a algo discreto. Procederemos a encontrar dicho conjunto de λ 's regresando a nuestro problema de optimización inicial y reescribiéndolo de manera adecuada, viz

$$-\mathbb{H}(f|f_0) + \sum_i \lambda_i \left(\int G_i(x) f(x) dx - C_i \right) \quad (\text{B.4})$$

usando (2.1), (2.2) y (2.3) en (2.4) tenemos que:

$$\begin{aligned} & - \int f \ln \left(\frac{f}{f_0} \right) + \sum_i \lambda_i \left(\int G_i f dx - C_i \right) = \\ & - \int f \left(\sum_i \lambda_i G_i + \mu - 1 \right) + \sum_i \lambda_i \left(\int G_i f dx - C_i \right) = \\ & - \int f \left(\sum_i \lambda_i G_i - \ln(Z(\lambda)) \right) + \sum_i \lambda_i \left(\int G_i f dx - C_i \right) = \\ & \int \left(- \sum_i f \lambda_i G_i + f \ln(Z(\lambda)) + \sum_i f \lambda_i G_i \right) - \sum_i \lambda_i C_i = \\ & \ln(Z(\lambda)) \int f dx - \lambda \cdot C = \\ & \ln(Z(\lambda)) - \lambda \cdot C \end{aligned}$$

Hemos llegado a una ecuación que sólo depende del vector λ , y como supusimos que f era óptimo entonces restaría resolver el problema de la minimización de la distancia entre $\ln(Z(\lambda))$ y $\lambda \cdot C$, o sea:

$$\inf_{\lambda} \{ \ln(Z(\lambda)) - \lambda \cdot C \}$$

La densidad que resuelve el problema de optimización es

$$f_{\lambda^*} = \frac{f_0 e^{\lambda^* \cdot G}}{Z(\lambda^*)}$$

donde λ^* es el punto donde $\ln(Z(\lambda)) - \lambda \cdot C$ alcanza el ínfimo.

Regularidad de la Solución

Haciendo uso de la fórmula hallada en la sección anterior demostraremos la regularidad de la solución.

Condiciones de Primer Orden

$$\begin{aligned} \frac{1}{Z(\lambda)} \cdot \frac{\partial Z(\lambda)}{\partial \lambda_j} - C_j &\Leftrightarrow \\ \frac{1}{Z(\lambda)} \cdot \frac{\partial Z(\lambda)}{\partial \lambda_j} &= C_j \end{aligned}$$

Condiciones de Segundo Orden

$$D_{\lambda_i \lambda_j}^2 (\ln(Z)) = \frac{Z_{\lambda_i \lambda_j} \cdot Z + Z_{\lambda_i} \cdot Z_{\lambda_j}}{Z^2} \quad (\text{C.1})$$

$$\begin{aligned} Z_{\lambda_i} &= \int f_0 G_{\lambda_i} e^{\lambda \cdot G} \\ Z_{\lambda_i \lambda_j} &= \int f_0 G_i G_j e^{\lambda \cdot G} \end{aligned}$$

Las ecuaciones anteriores en (2.5) dan como resultado que:

$$\begin{aligned} D_{\lambda_i \lambda_j}^2 (\ln(Z)) &= \frac{1}{Z} \int f_0 G_i G_j e^{\lambda \cdot G} - \frac{1}{Z} \int f_0 G_i e^{\lambda \cdot G} \cdot \frac{1}{Z} \int f_0 G_j e^{\lambda \cdot G} \\ &= \mathbb{E}^{f^\lambda} [G_i G_j] - \mathbb{E}^{f^\lambda} [G_i] \mathbb{E}^{f^\lambda} [G_j] \\ &= \text{Cov}^{f^\lambda} (G_i, G_j) = H_{i,j} \end{aligned}$$

De forma similar sea $h(X)$ el pago en el vencimiento de alguna opción y $\Pi(\lambda) = \mathbb{E}^{f^\lambda}[h(X)]$, luego:

$$\begin{aligned} \frac{\partial \Pi}{\partial \lambda_j} &= \frac{\partial}{\partial \lambda_j} \frac{\int f_0 e^{\lambda \cdot G} h dX}{\int f_0 e^{\lambda \cdot G} dX} \\ &= \text{Cov}^{f^\lambda}(h(X), G_j(X)) \\ \frac{\partial \Pi}{\partial C_j} &= \sum_i \frac{\partial \Pi}{\partial \lambda_i} \cdot \frac{\partial \lambda_i}{\partial C_j} \\ &= \sum_i \text{Cov}^{f^\lambda}(h(X), G_i(X)) \cdot H_{i,j}^{-1} \end{aligned}$$

Con lo cual Π como función de C_1, C_2, \dots, C_N es diferenciable y por lo tanto concluimos que la solución es regular, esto también es concluido en la referencia [2].

APÉNDICE D

Sobre la Aproximación de $\mathbb{H}(\mathbb{Q}; \mathbb{P})$ en un Espacio de Caminos

Considere dos medidas de probabilidad \mathbb{P} y \mathbb{Q} definidas en un espacio de caminos continuos $\Omega = \{S_t, 0 \leq t \leq T\}$ tales que

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu^{\mathbb{P}} dt + \sigma^{\mathbb{P}} dZ_t^{\mathbb{P}} \text{ bajo } \mathbb{P} \\ \frac{dS_t}{S_t} &= \mu^{\mathbb{Q}} dt + \sigma^{\mathbb{Q}} dZ_t^{\mathbb{Q}} \text{ bajo } \mathbb{Q} \end{aligned} \tag{D.1}$$

Donde $dZ^{\mathbb{P}}$ y $dZ^{\mathbb{Q}}$ son movimientos Brownianos con respecto a las medidas de probabilidad \mathbb{P} y \mathbb{Q} . De manera usual consideramos $t_0 = 0 < t_1, \dots < t_N = T$ una partición regular del intervalo $[0, T]$ con $\Delta t = t_j - t_{j-1}$ suficientemente pequeño.

Denote por $A_i^j(L, M) = j + i/2^L$, $-2^L \leq i \leq 2^L$, $-M \leq j \leq M$, para números enteros i, j, L, M .

Definimos

$$\pi_{n,L,M}^{\mathbb{P}}(i, j; k, l) = \begin{cases} \mathbb{P}\left(A_i^j \leq S_{t_{n+1}} < A_{i+1}^j \mid S_{t_{n-1}} = A_k^l\right) & \text{si } |j|, |l| \leq M, \\ 0 & M < |j|. \end{cases} \tag{D.2}$$

De la misma forma se puede definir $\pi_{n,L,M}^{\mathbb{Q}}(i, j; k, l)$.

Ahora consideramos la aproximación $\mathbb{H}^{\Delta t}(\mathbb{Q}; \mathbb{P})$ de la entropía relativa

$$\begin{aligned}
 \mathbb{H}^{\Delta t}(\mathbb{Q}; \mathbb{P}) &= \lim_{\substack{L \rightarrow \infty \\ M \rightarrow \infty}} \sum_{\{i_n, j_n; k_n, l_n\}} \ln \left(\frac{\prod_{n=0}^{N-1} \pi_n^{\mathbb{Q}}}{\prod_{n=0}^{N-1} \pi_n^{\mathbb{P}}} \right) \prod_{n=0}^{N-1} \pi_n^{\mathbb{Q}} \approx \mathbb{E}^{\mathbb{Q}} \left[\ln \left(\frac{\prod_{n=0}^{N-1} \pi_{n,\infty}^{\mathbb{Q}}}{\prod_{n=0}^{N-1} \pi_{n,\infty}^{\mathbb{P}}} \right) \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\ln \left(\prod_{n=0}^{N-1} \frac{\pi_{n,\infty}^{\mathbb{Q}}}{\pi_{n,\infty}^{\mathbb{P}}} \right) \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\sum_{n=0}^{N-1} \ln \left(\frac{\pi_{n,\infty}^{\mathbb{Q}}}{\pi_{n,\infty}^{\mathbb{P}}} \right) \right] \\
 &= \mathbb{E}^{\mathbb{Q}} \left[\sum_{n=0}^{N-1} \mathbb{E}^{\mathbb{Q}} \left[\ln \left(\frac{\pi_{n,\infty}^{\mathbb{Q}}}{\pi_{n,\infty}^{\mathbb{P}}} \right) \middle| S_n \right] \right], \tag{D.3}
 \end{aligned}$$

donde $\pi_{n,\infty}^{\mathbb{M}} = \pi_{n,L,M}^{\mathbb{M}}(i_n, j_n; k_n, l_n)$ con L, M suficientemente grandes y para $\mathbb{M} = \mathbb{P}, \mathbb{Q}$.

Bibliografía

- [1] Marco Avellaneda, Craig Friedman, Richard Holmes and Dominick Samperi. Calibrating volatility surfaces via relative-entropy minimization. *Applied Mathematical Finance*, 4:37-64, 1997.
- [2] Marco Avellaneda. Minimum-entropy calibration of asset-pricing models. *International Journal of Theoretical and Applied Finance*, 1, 1998.
- [3] E. Akyildirim, Y. Dolinsky, and H. Mete Soner. Approximating stochastic volatility by recombinaant trees. *ArXiv e-prints*, May 2012.
- [4] Per Christian Hansen. Rank-Deficient and Discrete Ill-Posed Problems, SIAM, 1997.
- [5] Bernt Øksendal. Stochastic Differential Equations, fifth edition. Springer-Verlag, 2003.
- [6] Nimalin Moodley. The Heston Model: A Practical Approach with MATLAB code, 2005.
- [7] Stephane Moins. Implementation of a Simulated Annealing algorithm for MATLAB, 2002.
- [8] Thomas M. Cover, Joy A. Thomas. *Elements of Information Theory*, Wiley-Interscience, 2006.
- [9] Jean-Pierre Fouque; George Papanicolaou; K. Ronnie Sircar. *Derivatives in Financial Markets with Stochastic Volatility*, Cambridge University Press, 2000.
- [10] S. Heston. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 1993 Volumen 6, número 2, pp.327-343.
- [11] Ioannis Karatzas, Steven E. Shreve. Brownian Motion and Stochastic Calculus, second edition, Springer, 1998.
- [12] Shunsuke Ihara. Information Theory for Continuous Systems, World Scientific, 1993.
- [13] Richard Aster, Brian Borchers, Clifford H. Thurber. Parameter Estimation and Inverse Problems, second edition, Academic Press, 2012.
- [14] Sergei Mikhailov, Ulrich Nöfel, 2003. *Heston's Stochastic Volatility Model Implementation, Calibration and Some Extensions*, Wilmott Magazine 4, pp.74-79.
- [15] Eric Benhamou, Emmanuel Gobet, Mohammed Miri. *Time dependent Heston model*. SIAM Journal on Financial Mathematics, SIAM, 2010, 1 (1), pp.289-325.

-
- [16] Patrick Billingsley, Probability and Measure, Wiley-Interscience, 1995.