Universidad Nacional Autónoma de México

Facultad de Ciencias

Riesgo de crédito

TESIS

que para obtener el título de:

ACTUARIO

presenta:

Edgar Díaz Ordoñez

Director de tesis: Mogens Bladt Petersen

2007

ÍNDICE GENERAL

Introducción

En el presente trabajo se examina al riesgo de crédito desde un punto de vista clásico. Se analizan supuestos económicos y algunas herramientas matemáticas para la elaboración de modelos de riesgo crédito.

En el primer capítulo se exponen los supuestos económicos y matemáticos para desarrollar un modelo en tiempo continuo. Se intenta utilizar sólo teoría de probabilidad elemental y cálculo para deducir los teoremas básicos requeridos para el análisis de modelos en tiempo continuo y, como parte de la deducción hacer explícitos los supuestos económicos implícitamente incluidos en los supuestos matemáticos.

En el segundo capítulo se pretende deducir un conjunto de restricciones para obtener una forma de valuación de opciones, esas restricciones son condiciones necesarias para que sean consistentes con la teoría del precio justo. También se introducen supuestos adicionales para examinar y extender la teoría del precio justo elaborada por Black-Scholes. Estas restricciones y supuestos están basadas en que las preferencias de los inversionistas tiende hacia una mejor posición siempre.

En el tercer capítulo se presenta una introducción a los conceptos de riesgo, riesgo de crédito; y el impacto que estos tienen en el mundo financieron. Además, se concluye con una una proposición la cual es una solución aproximada para la valuación de una opción de cambio que esta sujeta al riesgo de incumplimiento.

Por último, en el capítuo cuatro se dan las nociones de lo que es una cadana de Markov a tiempo continuo, su estructura, la matriz infinitesimal, entre otras cosa. En las secciones subsecuentes se presenta una breve introducción al concepto de calificaciones crediticias en la cual se describen los principales resultados sobre matrices de transición y se examinan diferentes métodos para estimar y comparar matrices de transición crediticias.

Supuestos económicos y matemáticos

Las herramientas matemáticas para la manipulación formal usada en el análisis de incertidumbre en tiempo continuo son algo especializadas y por consiguiente se puede no estar familiarizado. Además, en los libros concernientes al tema las deduciones son frecuentemente difíciles de seguir. Por otra parte, esas deduciones proporcionan una pequeña intuición en las relaciones entre los supuestos formales matemáticos y los correspondientes supuestos económicos. En este capítulo se intenta resolver el problema anterior utilizando sólo teoría de probabilidad elemental y cálculo para deducir los teoremas básicos requeridos para el análisis en tiempo continuo, y como parte de la deducción, hacer explícitos los supuestos económicos implícitos en los supuestos matemáticos. La siguiente exposición se sigue de Robert C. Merton [2].

Conceptos básicos

Definición 1 *Sea h el horizonte de intercambio que es la longitud mínima de tiempo en el que pueden hacerse sucesivas transacciones por los inversionistas en el mercado de valores.*

Se denotará a $X(t)$ como el precio de un valor al momento t, el cambio en el precio del valor entre el tiempo $t = 0$ y $T = nh > 0$ se podrá escribir como

$$X(T) - X(0) = \sum_{1}^{n} [X(k) - X(k-1)], \tag{1}$$

donde n es el número de intervalos comerciales entre el tiempo 0 y T y $X(k) - X(k-1)$ es la abreviación para $X(kh) - X[(k-1)h]$, que es el cambio del precio en el $k-$esimo intervalo comercial, $k = 1, 2, ..., n$.

Para deducir las implicaciones económicas del intercambio continuo, es necesario deducir las propiedades matemáticas de las series de tiempo de los cambios del precio en este entorno. Específicamente, las propiedades para el cambio del precio en un solo intervalo comercial y el cambio en un intervalo de tiempo finito fijo T cuando la longitud del intervalo comercial llega a ser muy pequeño y el número de intervalos comerciales n en $[0, T]$ se hace muy grande. Por consiguiente, debe señalarse que en ninguna parte en el análisis presentado aquí se asume que la distribución de $X(T) - X(0)$ es invariante a h.

Definición 2 *Sea el operador $\mathbb{E}_t$ la esperanza condicional en saber toda la información pertinente revelada a partir del tiempo t o antes de.*

Definición 3 *Sea $\epsilon(k)$ el cambio inesperado del precio en el valor entre $k-1$ y k, restringido a estar en el momento $k-1$, una variable aleatoria definida por*

$$\epsilon(k) \equiv X(k) - X(k-1) - I\!E_{k-1}\left\{X(k) - X(k-1)\right\}, \quad k = 1, ..., n., \qquad (2)$$

donde "el tiempo k" se usa como la abreviación para el "tiempo kh". Por construcción, $I\!E_{k-1}\{\epsilon(k)\} = 0$. Es más, por las propiedades de esperanza condicional, se tiene que $I\!E_{k-j}\{\epsilon(k)\} = 0$ para $j = 1, ..., k$. De la suma parcial $S_n \equiv \sum_1^n \epsilon(k)$ se forma una martingala [1]. No obstante, las propiedades de las martingalas de los rendimientos no esperados aquí es puramente un resultado de construcción, y por consiguiente no impone ningún supuesto económico. Sin embargo, dos supuestos económicos que se aplicarán son los siguientes.

Supuesto 1 *Para cada intervalo de tiempo finito $[0, T]$ existe un número $A_1 > 0$, independiente del número de intervalos comerciales n, tal que $var(S_n) \geq A_1$ donde la $var(S_n) \equiv I\!E_0\left\{[\sum_1^n \epsilon(k)]^2\right\}$.*

El supuesto 1 asegura que la incertidumbre asociada con los cambios no esperados del precio no se elimina incluso en el límite del intercambio continuo. Este supuesto es esencial para el modelo de intercambio continuo, esta propiedad fundamental se aprecia en el comportamiento del mercado de precios.

Supuesto 2 *Para cada intervalo de tiempo finito $[0, T]$, existe un número $A_2 < \infty$, independiente de n, tal que $var(S_n) \leq A_2$.*

El supuesto 2 asegura que la incertidumbre asociada con los cambios no esperados del precio en un periodo finito de tiempo no es tan grande que la variación pueda ser ilimitada.

Definición 4 *Sea $V(k) \equiv I\!E_0\left\{\epsilon^2(k)\right\}$, $k = 1, 2, ..., n$, la variación de las ganancias del valor entre el tiempo $k-1$ y k basándose en la información disponible a partir del tiempo cero, y sea $V \equiv max_k V(k)$.*

Supuesto 3 *Para cada intervalo de tiempo finito $[0, T]$, existe un número A_3, $1 \geq A_3 > 0$, independiente de n, tal que para $k = 1, ..., n$, $V(k)/V \geq A_3$.*

El supuesto 3 se relaciona estrechamente al supuesto 1 y en efecto queda fuera la posibilidad que toda la incertidumbre en los cambios no esperados del precio en $[0, T]$ se concentren en algún periodo pequeño de intercambio.

A estas alturas, se hará un breve paréntesis para definir algunos símbolos matemáticos que se usarán a lo largo del análisis en este capítulo.

Definición 5 *Sea $\psi(h)$ y $\phi(h)$ funciones de h. Se defininen los símbolos asintóticos $O[\phi(h)]$ y $o[\phi(h)]$ por*

[1] Para mayor información consultar Apéndice A.3.

$$\psi(h) = \begin{cases} O[\phi(h)] & \text{si} \quad \lim_{h\to 0}\left[\frac{\psi(h)}{\phi(h)}\right] \quad \text{es acotado,} \\ \\ o[\phi(h)] & \text{si} \quad \lim_{h\to 0}\left[\frac{\psi(h)}{\phi(h)}\right] = 0, \end{cases}$$

$$\psi(h) \sim {}^2\phi(h) \quad \text{si} \quad \psi(h) = O[\phi(h)] \;\; \text{y} \quad \psi(h) \neq o[\phi(h)].$$

En esencia, los símbolos de orden asintótico, $O(\cdot), o(\cdot)$, y $\sim$ se usarán para describir el comportamiento de la función $\psi(h)$ relativa a la función $\phi(h)$ para los valores de h cerca del cero.

Proposición 1 *Si los supuestos 1, 2, y 3 se aceptan, entonces $V(k) \sim h$, $k = 1, ..., n$. Es decir, $V(k) = O(h)$ y $V(k) \neq o(h)$, y $V(k)$ es asintóticamente proporcional a h donde el factor de proporcionalidad es positivo.*

Demostración

Sea

$$var(S_n) = I\!\!E_0 \left\{ \sum_1^n \sum_1^n \epsilon(k)\epsilon(j) \right\} = \sum_1^n \sum_1^n I\!\!E_0 \left\{ \epsilon(k)\epsilon(j) \right\}.$$

Considérese un término característico en la suma doble $I\!\!E_0 \{\epsilon(k)\epsilon(j)\}$. Supóngase $k \neq j$ y escójase $k > j$. Entonces,

$$I\!\!E_0 \{\epsilon(k)\epsilon(j)\} = I\!\!E_0 \{\epsilon(j) I\!\!E_j \{\epsilon(k)\}\}.$$

Pero, por construcción, $I\!\!E_j \{\epsilon(k)\} = 0$, $j < k$. Entonces, $I\!\!E_0 \{\epsilon(k)\epsilon(j)\} = 0$ para $k \neq j$. Por consiguiente, $var(S_n) = \sum_1^n V(k)$. De los supuestos 2 y 3, $nVA_3 \leq \sum_1^n V(k) \leq A_2$, y por consiguiente $V(k) \leq A_2 h / A_3 T$ donde $0 < A_2/A_3 < \infty$. Entonces, $V(k) = O(h)$. De los supuestos 3 y 1, $V(k) \geq A_1 A_3 h/T$ donde $A_1 A_3 > 0$. Por lo tanto, $V(k) \neq o(h)$. $\square$

Para algún intervalo comercial $[k-1, k]$, se asume que $\epsilon(k)$ puede tomar cualquiera de los m valores distintos denotados por $\epsilon_j(k)$, $j = 1, ..., m$ donde m es finito. Siempre que no haya ninguna ambigüedad sobre el tiempo k, se denotará $\epsilon_j(k)$ simplemente por ϵ_j. Supóngase además que existe un número $M < \infty$, independiente de n, tal que $\epsilon_j^2 \leq M$.

Si $p_j(k) \equiv I\!\!P \{\epsilon(k) = \epsilon_j | \text{la información disponible a partir del tiempo cero}\}$, entonces de la proposición 1 se sigue que

$$\sum_1^m p_j \epsilon_j^2 = O(h), \tag{3}$$

[2]Donde el símbolo $\sim$ significa "*es asintóticamente proporcional a*".

y puesto que m es finita se tiene de (3) que

$$p_j\epsilon_j^2 = O(h), \qquad j = 1, ..., m. \tag{4}$$

Cualquier evento j tal que $p_j\epsilon_j^2 = o(h)$ contribuirá asintóticamente una cantidad despreciable a la varianza de (1) porque $V(k) \neq o(h)$. Puesto que m es finito, se tiene que existen dos eventos por lo menos tal que $p_j\epsilon_j^2 \neq o(h)$, y de (4) $p_j\epsilon_j^2 \sim h$ para tales eventos. Ahora, sin pérdida de generalidad, se asume que $p_j\epsilon_j^2 \neq o(h)$, $j = 1, ..., m$, y por consiguiente $p_j\epsilon_j^2 \sim h$, $j = 1, 2, ..., m$.

Supuesto 4 *Para $j = 1, 2, ..., m$, p_j y ϵ_j son funciones suficientemente "bien-portadas" de h que existen números q_j y r_j tal que $p_j \sim h^{q_j}$ y $\epsilon_j \sim h^{r_j}$.*

Del supuesto 4, se tiene que $p_j\epsilon_j^2 \sim h^{q_j+2r_j}$. Pero $p_j\epsilon_j^2 \sim h$. De aquí, los valores asumidos por q_j y r_j no pueden ser arbitrarios, y de hecho deben satisfacer

$$q_j + 2r_j = 1 \qquad j = 1, 2, ..., m \tag{5}$$

La ecuación (5) define la relación entre estos dos números que deben satisfacer para cada evento j. Puesto que $p_j \leq 1$ y ϵ_j^2 esta acotado, ambos q_j y r_j deben ser no negativos, y por consiguiente, de (5), se tiene que $0 \leq q_j \leq 1$ y $0 \leq r_j \leq 1/2$, $j = 1, 2, ..., m$.

Como se mostrará, esos resultados localizados en los extremos del rango permisible para r_j determinarán las propiedades distribucionales asintóticas de $\epsilon(k)$. Será por consiguiente útil dividir los resultados en tres tipos: resultados del "tipo I" tal que $r_j = \frac{1}{2}$; resultados del "tipo II" tal que $0 < r_j < \frac{1}{2}$; y resultados del "tipo III" tal que $r_j = 0$.

Sea J la cual denotará el conjunto de eventos j tal que los resultados ϵ_j son del tipo I. Estos se siguen de (5) que, para $j \in J$, $q_j = 0$, y por consiguiente $p_j \neq o(1)$. Por otra parte, para todos los eventos $j \in J^c$ (es decir eventos del tipos II o III), se tiene que $p_j = o(1)$, y porque m es finito, la $\sum p_j = o(1)$, $j \in J^c$. Ahora, puesto que $\sum_1^m p_j = 1$, el conjunto J no puede estar vacío y, de hecho, implícitamente toda la densidad de probabilidad para $\epsilon(k)$ estará en eventos contenidos en J. En otros términos, para los intervalos comerciales pequeños h, implícitamente todas las observaciones de $\epsilon(k)$ serán del tipo I, y por consiguiente un nombre adecuado para J^c podría ser el "conjunto de eventos raros".

Trayectorias con "ningún evento raro"

En esta sección, se asumirá que todos los posibles resultados para $\epsilon(k)$, $k = 1, ..., n$, son del tipo I, y por consiguiente J^c está vacío, es decir, no hay ningún evento raro, y cada posible resultado ϵ_j, $j = 1, ..., m$, puede ocurrir.

Definición 6 *Sea α_k la esperanza condicional del rendimiento del peso por unidad de tiempo en un valor, definida por*

$$\alpha_k \equiv I\!\!E_{k-1}\left\{X(k) - X(k-1)\right\}/h \qquad k = 1, 2, ..., n \tag{6}$$

Supuesto 5 *Para cada h, se asume que α_k existe, $k = 1, ..., n$, y que existe un número $\alpha < \infty$, independiente de h, tal que $|\alpha_k| \leq \alpha$.*

El supuesto 5 simplemente asegura que para todos los valores con un precio finito la tasa esperada del rendimiento por unidad de tiempo en el horizonte de intercambio es finito, no importando que tan corto sea. Nótese que, no se asume que α_k sea constante con el tiempo, y de hecho α_k puede ser una variable aleatoria relacionada a la información disponible a partir de los primeros periodos $k-1$. De (2) y (6), puede escribirse el rendimiento del peso en un valor entre $k-1$ y k como

$$X(k) - X(k-1) = \alpha_k h + \epsilon(k) \qquad k = 1, 2, ..., n \tag{7}$$

Como se discutió en la sección anterior, un supuesto importante normalmente hecho en los modelos continuos es que la muestra de trayectorias de los precios de los valores son continuas.

Proposición 2 *Si para $k = 1, ..., n$, todos los posibles resultados para $\epsilon(k)$ son del tipo I, entonces la muestra de las trayectorias continuas para los precio de los valores serán continuas.*

Demostración

Sea $Q_k(\delta)$ la probabilidad que $|X(k) - X(k-1)| \geq \delta$ condicional de saber toda la información disponible a partir del periodo $k-1$. Una condición necesaria y suficiente para la continuidad[3] de la muestra de trayectorias para X es que, para cada $\delta > 0$, $Q_k(\delta) = o(h)$. Se define $\overline{u} = \text{máx}_{\{j\}} |\epsilon_j|/h^{1/2}$. Por hipótesis todos los resultados para $\epsilon(k)$ son tipo I, y por consiguiente $\overline{u} = O(1)$. Para cada número $\delta > 0$, se define la función $h^{+}(\delta)$ como la solución a la ecuación $\delta = \alpha h^{+} + \overline{u}(h^{+})^{1/2}$. Puesto que α y $\overline{u}$ son $O(1), h^{+}(\delta) > 0$ para cada $\delta > 0$. Es claro, que para toda $h < h^{+}(\delta)$ y cada posible resultado $X(k)$, $|X(k) - X(k-1)| < \delta$. Por lo tanto, para cada $h, 0 \leq h^{+}(\delta), Q_k(\delta) \equiv 0$, y el $\text{lím}[Q_k(\delta)/h] = 0$ cuando $h \to 0$ □

Como se ilustra en la Figura 1, mientras la muestra de trayectorias para $X(t)$ es continua, no es diferenciable en ningún punto casi seguramente. Por eso, el cálculo usual y la teoría estándar de ecuaciones diferenciales no puede usarse para describir la dinámica de movimientos del mercado de precios. Sin embargo, existe un cálculo generalizado y teoría correspondiente de ecuaciones diferenciales estocásticas que pueden usarse para estos casos.

En el desarrollo para la deducción de estos cálculos generalizados, será útil establecer ciertas propiedades de momento para $X(k) - X(k-1)$.

[3]Esta condición es llamada la “condición de Lindeberg”. Véase [**?**]

Definición 7 *Sea σ_k^2 la varianza condicional por unidad de tiempo del rendimiento del peso en el valor, definida por*

$$\sigma_k^2 \equiv I\!\!E_{k-1}\left\{\epsilon^2(k)\right\}/h \qquad k = 1, 2, ..., n \tag{8}$$

Puesto que para cada resultado ϵ_j, $\epsilon_j^2 = O(h)$, se tiene que $\sigma_k^2 = O(1)$. Por otra parte, de los supuestos 1 y 3, se tiene que $\sigma_k^2 > 0$ para toda h. Debido a que α_k esta acotada, se tiene que

$$I\!\!E_{k-1}\left\{[X(k) - X(k-1)]^2\right\} = \sigma_k^2 h + o(h) \qquad k = 1, 2, ..., n \tag{9}$$

Cuando h tiende a 0, los segundos momentos condicional central y no central de $X(k) - X(k-1)$ son los mismos. Nótese que no es supuesto que σ_k^2 es constante, y de hecho puede ser una variable aleatoria cuando examina periodos anteriores a $k-1$.

Considérese ahora el Nth momento absoluto incondicional de $\epsilon(k)$, para $2 < N < \infty$. Usando la misma definición para $\overline{u}$, se tiene que para $k = 1, 2, ..., n$

$$\begin{aligned} I\!\!E_0\left\{|\epsilon(k)|^N\right\} &= \sum_1^m p_j|\epsilon_j|^N \\ &\leq \sum_1^m p_j(\overline{u})^N h^{N/2} \\ &\leq \overline{u}^N h^{N/2} = o(h) \qquad \text{para} \quad N > 2 \end{aligned} \tag{10}$$

Así, todos los momentos absolutos de $\epsilon(k)$ superiores a el segundo son asintóticamente insignificantes de comparar con los primeros dos momentos. Similarmente,

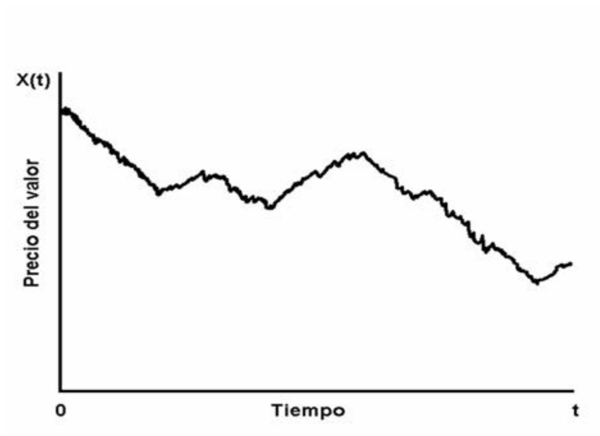

Figura 1: Muestra de trayectorias continuas para el precio de un valor con resultados "Tipo I"

se tiene que

$$\begin{aligned} \mathbb{E}_0\left\{|X(k)-X(k-1)|^N\right\} &\leq (\alpha h+\overline{u}h^{1/2})^N \\ &= \overline{u}^N h^{N/2}+o(h^{N/2}) \end{aligned} \tag{11}$$

Cuando $h^{N/2}$ tiende a 0, los Nth momentos absolutos incondicional central y no central de $X(k)-X(k-1)$ son el mismo.

Puesto que, las relaciones de orden entre los momentos descritos en (10) y (11) depende sólo en que $\{\epsilon_j\}=O(h^{1/2})$ y α_k sea acotada y no en las probabilidades de los resultados específicos $\{p_j\}$, se tiene inmediatamente que las relaciones de orden entre los momentos condicionales serán igual que para los momentos incondicionales. Por consiguiente,

$$\mathbb{E}_{k-1}\left\{|\epsilon(k)|^N\right\}=o(h) \quad \text{para} \;\; N\;>\;2 \tag{12}$$

y

$$\mathbb{E}_{k-1}\left\{|X(k)-X(k-1)|^N\right\}=\mathbb{E}_{k-1}\left\{|\epsilon(k)|^N\right\}+o(h^{N/2}) \tag{13}$$

Se define la variable aleatoria $u(k)$, $k=1,...,n$, por

$$u(k)\equiv \epsilon(k)/(\sigma_k^2 h)^{1/2} \tag{14}$$

donde, por construcción, $u_j \;\equiv\; \epsilon_j/(\sigma_k^2 h)^{1/2} \;=\; O(1)$, $\;j=1,...,m$; $\mathbb{E}_{k-1}\{u(k)\}=0$; $\mathbb{E}_{k-1}\left\{u^2(k)\right\}=1$; y $\mathbb{E}_{k-1}\left\{|u(k)|^N\right\}=O(1)$, $N>2$. Se puede reescribir a (7) como

$$X(k)-X(k-1)=\alpha_k h+\sigma_k u(k)h^{1/2} \qquad k=1,2,...,n \tag{15}$$

Ahora, siempre que los cambio del precio de un valor tengan sólo resultados tipo I, la dinámica del cambio de precio puede escribirse como una ecuación diferencial estocástica en forma de la ecuación (15) donde todas las variables aleatorias explícitas en el lado derecho son $O(1)$. Es más, a partir del tiempo $k-1$, la única variable aleatoria es $u(k)$, y en este caso (15) se llama ecuación diferencial estocástica condicional.

Habiendo establecido muchas de las propiedades asintóticas esenciales para $X(k)-X(k-1)$, ahora se discutirán las características de distribución de variables aleatorias que son función de los precios de los valores.

Definición 8 *Sea $F(t)$ una variable aleatoria dada por la siguiente regla $F(t)=f(X,t)$ si $X(t)=X$ donde f es una función C^2 con terceras derivadas parciales acotadas* [4]*.*

Siguiendo la convención establecida para $X(t)$, se usa la abreviación $F(k)$ para $F(kh)$ y $f[X(k),k]$ para $f[X(kh),kh]$. Supóngase que está en el momento

[4] El supuesto de que f tiene terceras parciales acotadas no es esencial, pero es hecho simplemente por conveniencia analítica.

$k-1$ y por consiguiente sabe los valores de $X(k-1),\ \alpha_k,\ \sigma_k$ y $\{p_j^{'}\}$ donde $p_j^{'}$ es la probabilidad condicional de que $u(k) = u_j,\ j = 1, ..., m$, dada la información disponible a partir del tiempo $k-1$.

Sea X el valor conocido de $X(k-1)$, entonces se definen los números $\{X_j\}$ por

$$X_j \equiv X + \alpha_k h + \sigma_k u_j h^{1/2} \qquad j = 1, 2, ..., m \tag{16}$$

Para cada valor X_j se puede usar el teorema de Taylor para escribir a $f(X_j, k)$ como

$$\begin{aligned} f(X_j, k) &= f(X, k-1) + f_1(X, k-1)(\alpha_k h + \sigma_k u_j h^{1/2}) \\ &\quad + f_2(X, k-1)h + \tfrac{1}{2} f_{11}(X, k-1)(\alpha_k h + \sigma_k u_j h^{1/2})^2 \\ &\quad + R_j \qquad\qquad j = 1, 2, ..., m \end{aligned} \tag{17}$$

donde los subíndices en f denotan las derivadas parciales y R_j se define por

$$\begin{aligned} R_j &\equiv \tfrac{1}{2} f_{22}(X, k-1)h^2 + f_{12}(X, k-1)(\alpha_k h + \sigma_k u_j h^{1/2})h \\ &\quad + \tfrac{1}{6} f_{111}(\eta_j, \varsigma_j)(X_j - X)^3 + \tfrac{1}{2} f_{112}(\eta_j, \varsigma_j)(X_j - X)^2 h \\ &\quad + \tfrac{1}{2} f_{122}(\eta_j, \varsigma_j)(X_j - X)h^2 + \tfrac{1}{6} f_{222}(\eta_j, \varsigma_j)h^3 \end{aligned} \tag{18}$$

donde $\eta_j \equiv X + \theta_j(X_j - X)$ y $\varsigma_j \equiv (k-1) + \nu_j$ para algún θ_j, ν_j tal que $0 \leq \theta_j \leq 1$ y $0 \leq \nu_j \leq 1$. Puesto que las terceras derivadas parciales de f son acotadas y $u_j = O(1),\ j = 1, ..., m$, se tiene por la sustitución para X_j de (1.16) en (1.18) que, para cada j,

$$|R_j| = O(h^{3/2}) = o(h) \qquad j = 1, 2, ..., m \tag{19}$$

Nótese que $(\alpha_k h + \sigma_k u_j h^{1/2})^2 = \sigma_k^2 u_j^2 h + o(h)$, entonces puede escribirse a (17) como

$$\begin{aligned} f(X_j, k) &= f(X, k-1) + f_1(X, k-1)(\alpha_k h + \sigma_k u_j h^{1/2}) \\ &\quad + f_2(X, k-1)h + \tfrac{1}{2} f_{11}(X, k-1)\sigma_k^2 u_j^2 h + o(h) \\ &\qquad\qquad j = 1, 2, ..., m \end{aligned} \tag{20}$$

Subsecuentemente (20) sostiene que para cada j, se puede describir la dinámica para $F(k)$ en forma de una ecuación diferencial estocástica condicional (aproximadamente) como

$$\begin{aligned} F(k) - F(k-1) &= \{f_1[X(k-1), k-1]\alpha_k + f_2[X(k-1), k-1] \\ &\quad + \tfrac{1}{2} f_{11}[X(k-1), k-1]\sigma_k^2 u^2(k)\big\} h \\ &\quad + f_1[X(k-1), k-1]\sigma_k u(k) h^{1/2} + o(h) \\ &\qquad\qquad k = 1, 2, ..., n \end{aligned} \tag{21}$$

donde (21) es condicionada en saber $X(k-1),\ \alpha_k$ y σ_k.

Definición 9 *Sea* $\mu_k \equiv I\!E_{k-1}\left[F(k) - F(k-1)\right]/h$ *la esperanza condicional del cambio en* F *por unidad de tiempo*

Aplicando formalmente el operador $I\!\!E_{k-1}$ esperanza condicional a ambos lados de (21) y sumando entonces de $j = 1, ..., m$. Nótese que las derivadas de f en el lado derecho de (21) son evalúadas en $X(k-1)$ y son por consiguiente no estocásticas relativas al tiempo $k-1$, se tiene que, para $k = 1, ..., n$,

$$\begin{aligned} \mu_k &= \{f_1[X(k-1), k-1]\alpha_k + f_2[X(k-1), k-1] \\ &\quad + \tfrac{1}{2} f_{11}[X(k-1), k-1]\sigma_k^2\} + o(1) \end{aligned} \tag{22}$$

Como $\alpha_k,\ \mu_k = O(1)$, y a ese orden es completamente determinado conocer $X(k-1)$ y sólo los primeros dos momentos para el cambio en X.

Sustituyendo de (22), se puede reescribir a (21) como

$$\begin{aligned} F(k) - F(k-1) &= \mu_k h \ + \ \tfrac{1}{2} f_{11}[X(k-1), k-1]\sigma_k^2[u^2(k) - 1]h \\ &\quad + \ f_1[X(k-1), k-1]\sigma_k u(k) h^{1/2} \ + \ o(h) \end{aligned} \tag{23}$$

El análisis de (23) muestra que, al tender h, la ecuación diferencial estocástica condicional para $F(k) - F(k-1)$ es esencialmente de la misma forma como la ecuación (15) para $X(k) - X(k-1)$ salvo la componente estocástica $O(h)$ adicional. De hecho, de (23), se pueden escribir los momentos condicionales para $F(k) - F(k-1)$ como

$$I\!\!E_{k-1}\left\{[F(k) - F(k-1)]^2\right\} = \{f_1[X(k-1), k-1]\sigma_k\}^2\, h + o(h) \tag{24}$$

y

$$I\!\!E_{k-1}\left\{[F(k) - F(k-1)]^N\right\} = O(h^{N/2}) = o(h) \qquad N > 2 \tag{25}$$

De la relación de orden para los momentos condicionales de $F(k) - F(k-1)$ es igual que para los momentos condicionales de $X(k) - X(k-1)$, y $O(h)$ el componente estocástico hace una contribución insignificante a los momentos de $F(k) - F(k-1)$.

De hecho, no sólo es la relación de orden de los propios momentos para los cambios en F y X el mismo, sino los comomentos entre ellos de cambios contemporáneos tienen la misma relación de orden. Es decir, de (15) y (23) se tiene que

$$\begin{aligned} &I\!\!E_{k-1}\{[F(k) - F(k-1)][X(k) - X(k-1)]\} \\ &\quad = \left\{f_1[X(k-1), k-1]\sigma_k^2\right\} h \ + \ o(h) \end{aligned} \tag{26}$$

y

$$\begin{aligned} I\!\!E_{k-1}\left\{[F(k) - F(k-1)]^j[X(k) - X(k-1)]^{N-j}\right\} &= O(h^{N/2}) = o(h) \\ j = 1, ..., N \qquad N > 2 & \end{aligned} \tag{27}$$

Aunque (26) y (27) son los comomentos no centrales, la diferencia entre los comomentos central y no central será $o(h)$, y por consiguiente los dos pueden usarse indistintamente.

Finalmente, se tiene un resultado muy poderoso que, al tender h, los cambios simultáneos en F y X son perfectamente correlacionados. Así, si ρ_k se define como el coeficiente de correlación condicional por unidad de tiempo entre los cambios simultáneos en F y X, entonces, de (24) y (26),

$$\rho_k = \begin{cases} 1 \ + \ o(1) & \text{si } f_1[X(k-1), k-1] > 0 \\ -1 \ + \ o(1) & \text{si } f_1[X(k-1), k-1] < 0 \end{cases}$$

Así, incluso si F es una función no lineal de X, en el límite sus cambios instantáneos serán perfectamente correlacionados.

Habiéndose demostrado que el término estocástico $O(h)$ contribuye una cantidad despreciable a la varianza en F en un intervalo muy corto, se estudiará su contribución ahora al cambio en F en uno finito, y no necesariamente pequeño, intervalo. Se define la variable aleatoria $G(t)$ por

$$\begin{aligned} G(k) - G(k-1) \quad \equiv \quad & F(k) - F(k-1) - \mu_k h \\ & -f_1[X(k-1), k-1]\sigma_k u(k) h^{1/2} \\ & \qquad\qquad k = 1, ..., n \end{aligned} \tag{28}$$

Aquí, $G(k) - G(k-1)$ es el error de aproximar $F(k) - F(k-1)$ por $\mu_k h + f_1 \sigma_k u(k) h^{1/2}$. Si se define

$$y(k) \equiv \tfrac{1}{2} f_{11}[X(k-1), k-1]\sigma_k^2[u^2(k) - 1]$$

entonces de (23) se puede reescribir a (28) como

$$G(k) - G(k-1) = y(k)h + o(h) \qquad k = 1, ..., n \tag{29}$$

Por construcción, $\mathbb{E}_{k-1}\{y(k)\} = 0$, y por consiguiente $\mathbb{E}_{k-j}[y(k)] = 0, \ j = 1..., k$. Por consiguiente, la suma parcial $\sum_1^n y(k)$ forma una martingala. Debido a que $\mathbb{E}_0 \left\{\sum_1^n y^2(k)/k^2\right\} < \infty$ cuando $n \to \infty$, se sigue de la Ley de los Grandes Números para las martingalas que

$$\lim_{h \to \infty} \left[h \sum_1^n y(k) \right] = T \lim_{h \to \infty} \left[\frac{1}{n} \sum_1^n y(k) \right] \to 0 \tag{30}$$

De (29), se tiene que para T ($\equiv nh$) > 0 fija

$$\begin{aligned} G(T) - G(0) \quad &= \quad h \sum_1^n y(k) + \sum_1^n o(h) \\ &= \quad h \sum_1^n y(k) + o(1) \end{aligned} \tag{31}$$

Tomando el límite de (31) cuando $n \to \infty$ ($h \to 0$), se tiene de (30) que $G(T) - G(0) \to 0$. Es decir, el error acumulativo de la aproximación tiende a cero con probabilidad uno.

Ahora, para $T > 0$, se tiene de (28) que en el límite (cuando $h \to 0$)

$$\begin{aligned} F(T) - F(0) &= \sum_1^n [F(k) - F(k-1)] \\ &= \sum_1^n \mu_k h + \sum_1^n f_1[X(k-1), k-1]\sigma_k u(k) h^{1/2} \end{aligned} \tag{32}$$

con probabilidad uno. Ahora, en el límite, el término estocástico $O(h)$ en (23) tendrá un efecto despreciable en el cambio en F en un intervalo finito.

Es natural interpretar el límite de las sumas en (32) como integrales. Para cada k, $k = 1, ..., n$, se define $t \equiv kh$. Por los usuales argumentos para la integración de Riemann, se tiene que

$$\lim_{n\to\infty} \left(\sum_1^n \mu_k h \right) = \int_0^T \mu(t) dt \tag{33}$$

donde $\mu(t)$ es el límite de μ_k y es llamado la esperanza condicional instantánea del cambio en F por unidad de tiempo, condicional en la información disponible al momento t. Claro, debido al coeficiente de $h^{1/2}$, la segunda suma no satisfará las condiciones usuales para la integral de Riemann. Sin embargo, se puede proceder formalmente y definir la integral estocástica como el límite de la suma dada por

$$\lim_{n\to\infty} \left\{ \sum_1^n f_1[X(k-1), k-1]\sigma_k u(k) h^{1/2} \right\} = \int_0^T f_1[X(t), t]\sigma(t) u(t) (dt)^{1/2} \tag{34}$$

donde el formalismo $(dt)^{1/2}$ se usa para distinguir esta integral de la integral usual de Riemann en (33). Ahora, se tiene de (32) que el cambio en F entre 0 y T puede escribirse

$$F(T) - F(0) = \int_0^T \mu(t) dt + \int_0^T f_1[X(t), t]\sigma(t) u(t) (dt)^{1/2} \tag{35}$$

Dada esta representación de la integral estocástica para el cambio en F en un intervalo de tiempo, se procede formalmente a definir el diferencial estocástico para F por

$$dF(t) = \mu(t) dt + f_1[X(t), t]\sigma(t) u(t) (dt)^{1/2} \tag{36}$$

donde la forma diferencial dF se usa en lugar de la notación usual, dF/dt, para subrayar el resultado previamente discutido que las muestras de trayectorias no son casi en ninguna parte diferenciables en el sentido usual.

La correspondiente representación de la integral estocástica y diferencial para la dinámica de $X(t)$ puede escribirse inmediatamente de (35) y (36) simplemente escogiendo $f(X, t) = X$ esto es, de (35),

$$X(T) - X(0) = \int_0^T \alpha(t) dt + \int_0^T \sigma(t) u(t) (dt)^{1/2} \tag{37}$$

y

$$dX(t) = \alpha(t)dt + \sigma(t)u(t)(dt)^{1/2} \tag{38}$$

donde en este caso el término estocástico omitido $O(dt)$ es idénticamente cero porque $f_{11} \equiv 0$.

Las únicas restricciones en la distribución para $u(t)$ fueron:

a) $I\!E\{u(t)\} = 0$,

b) $I\!E\{u^2(t)\} = 1$,

c) $u(t) = O(1)$,

d) y que la distribución de $u(t)$ sea discreta.

Las restricciones (a) y (b) están puramente por construcción, y (c) y (d) pueden relajarse para permitir a la mayoría de las distribuciones continuas "bien-portadas" incluyendo algunas con dominio no acotado. En particular, no fue supuesto que $\{u(t)\}$ se distribuya idénticamente o consecutivamente independiente. Sin embargo, desarrollar el análisis más allá requiere un supuesto económico adicional.

Supuesto 6 *El proceso estocástico para $X(t)$ es un proceso de Markov. En otros términos, la distribución de probabilidad condicional para los valores futuros de X, condicional en estar en el momento t, sólo depende del valor actual de X y la inclusión de información extensa disponible a partir de esa fecha no alterará esta probabilidad condicional.*

Mientras este supuesto puede parecer ser bastante restrictivo, pueden transformarse muchos procesos que no son formalmente Markov a la forma de Markov por el método de "expansión de los estados" [5] y por consiguiente el supuesto 6 podría relajarse para decir que las probabilidades condicionales para X dependen en sólo una cantidad finita de información del pasado. Del supuesto 6 puede escribirse la densidad de probabilidad condicional para $X(T) = X$ en el momento T, condicional en $X(t) = x$, como

$$p(x,t) \equiv p(x,t;X,T) = I\!P\{X(T) = X | X(t) = x\} \qquad t < T \tag{39}$$

donde la omisión de los argumentos explícitos X y T se entenderán para significar la pertenencia de estos dos valores fijos. Ahora, para X y T fijos, $p[X(t),t]$ (visto antes del tiempo t) es una variable aleatoria que es una función del precio del valor al momento t. Por consiguiente, con tal de que p sea una función bien-portada de x y t, satisfará todas las propiedades previamente deducidas para $F(t)$. En particular, dp satisfará (36) donde $\mu(t)$ es la esperanza condicional del

[5] Para una mejor discusión y extensas referencias del método consultar Cox and Miller (1968)

cambio por unidad de tiempo en p. Sin embargo, p es una densidad de probabilidad, y por consiguiente su cambio *esperado* es cero. Tomando el límite de (22) cuando $h \to 0$ y aplicando la condición que $\mu(t) = 0$, se tiene que

$$0 = \frac{1}{2}\sigma^2(x,t)p_{11}(x,t) \ + \ \alpha(x,t)p_1(x,t) \ + \ p_2(x,t) \tag{40}$$

donde los subíndices en p denotan las derivadas parciales. Es más, por el supuesto 6, $\alpha(t)$ y $\sigma(t)$ son, a lo más, funciones de $x(t)$ y t. Aquí, se hace esta dependencia explícita volviendo a escribir estas funciones como $\alpha(x,t)$ y $\sigma^2(x,t)$ respectivamente. Un análisis de (40) mustra que es una ecuación diferencial particial lineal del tipo parabólico y es a veces llamada la ecuación “backward de Kolmogorov”[6]. Por consiguiente, sujeta a condiciones de frontera, (40) completamente especifica las densidades de probabilidad de transición para el precio del valor. Así, en el límite, el conocimiento de las dos funciones $\sigma^2(x,t)$ y $\alpha(x,t)$ es suficiente para determinar la distribución de probabilidad para el cambio en el precio de un valor entre cualquiera dos fechas.

Por consiguiente, se sigue que las únicas características de las distribuciones para $\{u(t)\}$ que afectan la distribución asintótica para el precio del valor son el primero y segundo momento, y, por construcción, ellos son constantes a través del tiempo. Es decir, salvo el requisito de ajuste en los primeros dos momentos, las características distribucionales de $\{u(t)\}$ pueden escogerse casi arbitrariamente sin tener cualquier efecto en la distribución asintótica para el precio del valor. Ahora, en el límite, ningún contenido económico está perdido asumiendo que $\{u(t)\}$ son independientes e idénticamente distribuídos, y por consiguiente para el resto de esta sección se hace este supuesto. Advertencia: este supuesto no implica que el cambio en $X(t)$ o $F(t)$ tiene estas propiedades. De hecho, si cualquiera $\alpha(t)$ o $\sigma^2(t)$ es una función de $X(t)$, entonces los cambios en $X(t)$ no serán independientes ni idénticamente distribuidos.

Definición 10 *Sea $Z(t)$ una variable aleatoria cuyo cambio de valor en el tiempo se describe por una ecuación diferencial estocástica como (15) pero con $\alpha_k \equiv 0$ y $\sigma_k \equiv 1$, $k = 1, ..., n$.*

En otros términos, la esperanza condicional del cambio en $Z(t)$ por unidad de tiempo es cero y la varianza condicional de ese cambio por unidad de tiempo es uno. Por consiguiente,

$$\begin{aligned} Z(T) - Z(0) &= \sum_1^n [Z(k) - Z(k-1)] \\ &= h^{1/2} \sum_1^n u(k) \\ &= T^{1/2} \frac{\sum_1^n u(k)}{n^{1/2}} \end{aligned} \tag{41}$$

[6]Ver Cox and Miller (1968, pág. 215)

Las $\{u(k)\}$ son independientes e idénticamente distribuidas con media igual a cero y varianza igual a uno. Por consiguiente, por el Teorema del Límite Central, $\left\{\sum_1^n u(k)/n^{1/2}\right\}$ tendrá una distribución normal estándar. Así se sigue de (41) que, $Z(T) - Z(0)$ se distribuirá normal con media cero y una varianza igual a T para todo $T > 0$. De hecho, la solución a (40) con $\sigma^2 = 1$ y $\alpha = 0$ es

$$p(x, t; X, T) = \frac{\exp[-(X - x)^2/2(T - t)]}{[2\pi(T - t)]^{1/2}} \tag{42}$$

que es una función de densidad normal.

Ya que la distribucionalidad escogida para $\{u(t)\}$ puede hacerse casi arbitrariamente y la distribución para $Z(T) - Z(0)$ es Gaussiana para toda T finita, es natural y conveniente asumir que $\{u(t)\}$ son distribuidas normal estándar. En una forma análoga a (38), se puede escribir la representación de la ecuación diferencial estocástica para $Z(t)$ como

$$dZ(t) = u(t)(dt)^{1/2} \tag{43}$$

En el caso donde las $\{u(t)\}$ sean independientes y se distribuyen normal estándar, el proceso dZ descrito en (43) se llama proceso de Wiener o Movimiento Browniano, y se reservará la notación dZ para denotar tal proceso a lo largo de este capítulo.

Puesto que la distribución seleccionada para $\{u(t)\}$ no afecta la distribución para X, sin pérdida de generalidad puede escribirse la representación de la integral estocástica y del diferencial para la dinámica de $X(t)$, de las ecuaciones (37) y (38), como

$$X(T) - X(0) = \int_0^T \alpha(t)X(t)dt + \int_0^T \sigma(t)X(t)dZ(t) \tag{44}$$

y

$$dX(t) = \alpha(t)X(t)dt + \sigma(t)X(t)dZ(t) \tag{45}$$

Se sigue inmediatamente de (35) y (36) que, si la dinámica de $X(t)$ puede describirse por un proceso de Itô, entonces la dinámica de funciones bien portadas de $X(t)$ también se representará por un proceso de Itô. Esta relación entre la dinámica de $X(t)$ y $F(t)$ se formaliza en el lema siguiente.

Lema 1 (Lema de Itô) *Sea $f(X, t) : \mathbb{R} \to [0, \infty]$ una función C^2 y tomando la integral estocástico definida por (44), entonces la variable aleatoria $F \equiv f$ es una integral estocástica y su diferencial estocástico es*

$$dF = f_1(X, t)dX + f_2(X, t)dt + \frac{1}{2}\left[f_{11}(X, t) + f_{12}(X, t) + f_{22}(X, t)\right](dX)^2$$

Como $dZ = u(t)(dt)^{1/2}$ es un proceso de Wiener. Nótese que $E(dZ) = 0$ y $\text{var}(dZ) = dt$, también sería interesante saber la conducta de $(dZ)^2$ y $dt\,dZ$. La media y varianza de $(dZ)^2$ y $dt\,dZ$ se demostrarán con poco rigor matemático.

$$E((dZ)^2) = \text{var}(dZ) + [E(dZ)]^2 = dt$$

$$\mathrm{var}((dZ)^2) = E([dZ]^4) - \left[E(dZ)^2\right]^2 = o(dt)$$

$$E(dt\, dZ) = E(dt[u(t)(dt)^{1/2}]) = 0$$

$$\mathrm{var}(dt\, dZ) = E(dt^2\, [dZ]^2) - [E\,(dt(dZ))]^2 = o(dt)$$

Supóngase que se tratan los términos de orden $o(dt)$ como esencialmente cero, se observa que $(dZ)^2$ y $dtdZ$ son ambos no estocásticos, ya que su varianza es esencialmente cero. Por lo tanto, $(dZ)^2 = dt$ y $dtdZ = 0$.

Todos los resultados deducidos en esta sección son basado en el supuesto que los cambios en el precio de los valores son todos resultados tipo I . Esta clase de procesos es sólo un subconjunto del conjunto de procesos que satisfacen los supuestos económicos 1-6. Ahora, para completar el estudio de los modelos matemáticos de intercambio continuo, se proporcionará un análisis de esos procesos que permiten la posibilidad de "eventos raros".

Trayectorias con "eventos raros"

En esta sección, es supuesto que los resultados para $\epsilon(k),\ k = 1, ..., n$, puede ser cualquier resultado tipo I o tipo II, pero no el tipo III. Así, permite la posibilidad de eventos raros con resultados tipo II aunque, como se mostró en la sección anterior, implícitamente todas las observaciones de $\epsilon(k)$ serán del tipo I.

La forma del análisis presentado aquí es esencialmente igual que en la sección anterior. De hecho, la conclusión principal de este análisis será que, las propiedades de la distribución de los rendimientos de un valor son indistinguibles de aquéllas de la sección 1.1, en otros términos, eventos raros con resultados tipo II "no importan".

Para mostrar esto, primero se demostrará que, la muestra de trayectorias para los precios de un valor es continua con el tiempo.

Definición 11 *Para cada periodo de tiempo k, se define $r \equiv \min r_j$ donde el comportamiento del término de orden para ϵ_j es $h^{r_j},\ j = 1, ..., m$.*

Puesto que todos los resultados son tipo I o tipo II, $r > 0$, y $|\epsilon_j| = O(h^r),\ j = 1, ..., m$.

Proposición 3 *Si para $k = 1, ..., n$, todos los posibles resultados para $\epsilon(k)$ son cualquier resultado tipo I o tipo II, entonces la trayectoria continua del precio de un valor será continua.*

Demostración

Sea $Q_k(\delta)$ la probabilidad que $|X(k) - X(k-1)| \geq \delta$ condicional en saber toda la información disponible a partir del tiempo $k-1$. Como en la demostración de la proposición 2, un requisito y condición suficiente para la continuidad de la muestra de trayectorias para X es que, para cada $\delta > 0$, $Q_k(\delta) = o(h)$. Se define a $\overline{u} \equiv \text{máx}_{\{j\}} |\epsilon_j| h^{-r}$. Por definición de r, $\overline{u} = O(1)$. Puesto que cada número $\delta > 0$, se define la función $h^+(\delta)$ como la solución a la ecuación $\delta = \alpha h^+ + \overline{u}(h^+)^r$ donde, por el supuesto 5, α es $O(1)$. Puesto que $r > 0$ y α y $\overline{u}$ son ambos $O(1)$, existe una solución $h^+(\delta) > 0$ para cada $\delta > 0$. Por consiguiente, para todo $h < h^+(\delta)$ y cada posible resultado $X(k)$, $|X(k) - X(k-1)| < \delta$. Por lo tanto, para cada h, $0 \leq h < h^+(\delta)$, $Q_k(\delta) \equiv 0$, y $\text{lím}[Q_k(\delta)/h] = 0$ cuando $h \to 0$ $\square$

Habiéndose establecido la continuidad de la muestra de la trayectoria, se mostrará ahora que las propiedades de momento para $X(k) - X(k-1)$ son iguales que en la sección 1.1. Del supuesto 5, $E_{k-1}\{X(k) - X(k-1)\}$ es asintóticamente proporcional a h, y por consiguiente lo es también para $E_0\{X(k) - X(k-1)\}$. As, de la proposición 1 y la ecuación (5), la varianza incondicional de $X(k) - X(k-1)$ es asintóticamente proporcional a h. El Nth momento absoluto incondicional de $\epsilon(k)$, $2 < N < \infty$, puede escribirse como

$$
\begin{aligned}
E_0\left\{|\epsilon(k)|^N\right\} &= \sum_1^m p_j |\epsilon_j|^N \\
&= O\left(\sum_1^m h^{(N-2)r_j+1}\right) \\
&= O\left(h^{(N-2)r+1}\right) \\
&= o(h) \qquad N > 2
\end{aligned}
\tag{46}
$$

puesto que $r > 0$. Así, todos los momentos absolutos de $\epsilon(k)$ superiores que el segundo son asintóticamente insignificantes en comparación con los primeros dos momentos. Es más, por el supuesto 5, estas mismas relaciones de orden se obtendrán para ambos momentos central y no central de $X(k) - X(k-1)$.

Con tal de que las relaciones de orden entre las probabilidades incondicional y condicional permanezcan igual, los momentos condicionales de $X(k) - X(k-1)$ tiene las mismas propiedades de orden como los momentos incondicionales: es decir

$$E_{k-1}\left\{[X(k) - X(k-1)]^2\right\} = \sigma_k^2 h \; + \; o(h) \qquad k = 1, ..., n \tag{47}$$

donde σ_k^2 es la varianza condicional por unidad de tiempo definida en (8) y $\sigma_k^2 > 0$ y $O(1)$, y

$$E_{k-1}\left\{|X(k) - X(k-1)|^N\right\} = o(h) \qquad \text{para } N > 2 \tag{48}$$

Aquí, las relaciones de momento para $X(k) - X(k-1)$ son idénticas con aquellas deducidas en la sección 1.1 donde sólo resultados tipo I se permitieron.

Para completar el análisis, se examinarán las características de la distribución de variables aleatorias que son funciones de los precios de los valores. Sea $F(t)$ una variable aleatoria dada por la regla que $F(t) = f(X)$ si $X(t) = X$. Nótese que, a diferencia del análisis paralelo en la sección 1.1, la dependencia explícita de f en t se ha eliminado. Esto se hace solamente para mantener la notación y simplificar relativemente el análisis. Sin embargo, incluso la dependencia explícita del tiempo no cambiaría el método de deducción o las conclusiones.

Se define a K para ser el entero más pequeño tal que $Kr \geq 1$. Puesto que $r > 0$, K es finito. Si f es una función C^2 con derivada de orden $(K+1)th$ acotada, entonces del teorema de Taylor y (47) y (48) se tiene que

$$I\!E_{k-1}\left\{F(k) - F(k-1)\right\} = \left\{f^{(1)}[X(k-1)]\alpha_k + \frac{1}{2}f^{(2)}[X(k-1)]\sigma_k^2\right\} h + o(h) \tag{49}$$

donde $f^{(1)}$ denota la ith derivada de f.

Es más, es claro mostrar que los momentos condicionales para $F(k)-F(k-1)$ aquí son igual que los deducidos en las ecuaciones (24) y (25) de la sección 1.1; es decir,

$$I\!E_{k-1}\left\{[F(k) - F(k-1)]^2\right\} = \left\{f^{(1)}[X(k-1)]\sigma_k\right\}^2 h + o(h) \tag{50}$$

y

$$I\!E_{k-1}\left\{[F(k) - F(k-1)]^N\right\} = o(h) \qquad \text{para } N > 2 \tag{51}$$

Ahora, la relación de orden para los momentos condicionales de $F(k)-F(k-1)$ es igual que para los momentos condicionales de $X(k)-X(k-1)$. Por consiguiente, en los intervalos cortos, la parte inesperada del cambio en F aquí será dominada por el término $f^{(1)}[X(k-1)]\epsilon(k)$ en la misma forma que se denominó el cambio en F en la sección 1.1.

Habiéndose estudiado el cambio en F en un intervalo de tiempo muy corto, se examinarán las propiedades estocásticas ahora para el cambio en F en uno finito, y no necesariamente pequeño, intervalo de tiempo. Para cada k, $k = 1, ..., n$, se definen las variables aleatorias $\{y_j(k)\}$ como

$$\begin{aligned} y_j(k) \quad &\equiv \quad f^{(j)}[X(k-1)]\,\frac{[X(k)-X(k-1)]^j - I\!E_{k-1}\left\{[X(k)-X(k-1)]^j\right\}}{j!} \\ &\qquad\qquad j = 2, ..., K \end{aligned} \tag{52}$$

Aún más se define la variable aleatoria $G(t)$ por

$$\begin{aligned} G(k) - G(k-1) \quad &\equiv \quad F(k) - F(k-1) - I\!E_{k-1}\left\{F(k) - F(k-1)\right\} \\ &\qquad - f^{(1)}[X(k-1)]\epsilon(k) \qquad k = 1, ..., n \end{aligned} \tag{53}$$

que por el teorema de Taylor puede escribirse como

$$G(k) - G(k-1) = \sum_{j=2}^{K} y_j(k) \; + \; R_{K+1} \tag{54}$$

donde R_{K+1} se define por

$$\begin{aligned} R_{K+1} \quad &\equiv \quad f^{(K+1)}[\theta X(k-1) + (1-\theta)X(k)] \\ &\qquad \times \frac{[X(k)-X(k-1)]^{K+1} - I\!E_{k-1}\left\{[X(k)-X(k-1)]^{K+1}\right\}}{(K+1)!} \end{aligned} \tag{55}$$

para algún $\theta,\ 0 \leq \theta \leq 1$. Pero $f^{(K+1)}$ esta acotada y

$$[\alpha_k h + \epsilon(k)]^{K+1} = O(h^{r(K+1)})$$

para cada posible resultado para $\epsilon(k)$. Ahora, puesto que $rK \geq 1,\ R_{K+1} = o(h)$ Por consiguiente, se pouede volver a escribir (54) como

$$G(k) - G(k-1) = \sum_{j=2}^{k} y_j(k) + o(h) \qquad k = 1, ..., n \tag{56}$$

De (56), se puede escribir la varianza incondicional de $G(k) - G(k-1)$ como

$$\begin{aligned}
\text{var}[G(k) - G(k-1)] &= I\!E_0 \left\{ \sum_2^K \sum_2^K y_i(k) y_j(k) \right\} + o(h) \\
&\leq \sum_2^K \sum_2^K \ M_i M_j I\!E_0 | \left\{ [X(k) - X(k-1)]^i \right. \\
&\quad \left. - \ I\!E_{k-1}[X(k) - X(k-1)]^i \right\} \left\{ [X(k) - X(k-1)]^j \right. \\
&\quad \left. - \ I\!E_{k-1}[X(k) - X(k-1)]^j \right\} | / i! j! + o(h)
\end{aligned} \tag{57}$$

donde M_i es la mínima cota superior en $|f^{(i)}|,\ i = 2, ..., K$. De (46) y (57), se tiene, por consiguiente, que

$$\text{var}[G(k) - G(k-1)] = O(h^{2r+1}) \ + \ o(h) = o(h) \tag{58}$$

porque $r > 0$. Para el intervalo finito $[0, T]$, se tiene que

$$\begin{aligned}
G(T) - G(0) &= \sum_{k=1}^{n} [G(k) - G(k-1)] \\
&= \sum_1^n \sum_2^K y_j(k) + o(1)
\end{aligned} \tag{59}$$

$\sum_1^n \sum_2^K y_j(k)$ forma una martingala, y por consiguiente la varianza incondicional de $G(T) - G(0)$ puede escribirse como

$$\text{var}[G(T) - G(0)] = \sum_1^n \text{var}[G(k) - G(k-1)] + o(1) \tag{60}$$

Pero, de (58) y (60), se sigue que

$$\text{var}[G(T) - G(0)] = O(h^{2r}) + o(1) = o(1) \tag{61}$$

y por consiguiente, la varianza de $G(T) - G(0)$ tiende a cero para cada intervalo finito $[0, T]$.

Ahora, para $T > 0$, se tiene de (49), (50), (51), y (61) que

$$\begin{aligned} F(T) - F(0) &= \sum_1^n [F(k) - F(k-1)] \\ &= \sum_1^n \mu_k h + \sum_1^n f^{(1)}[X(k-1)]\epsilon(k) \end{aligned} \tag{62}$$

donde μ_k es la esperanza condicional del cambio en F por unidad de tiempo definido en (22). En una forma similar al análisis en la sección 1.1, se puede expresar formalmente el límite de las sumas (62) como la suma de dos integrales, es decir,

$$F(T) - F(0) = \int_0^T \mu(t)dt + \int_0^T f^{(1)}[X(t)]\epsilon(t) \tag{63}$$

como en (36) de la sección 1.1, se puede definir el diferencial estocástico formalmente para F por

$$dF(t) = \mu(t)dt \ + \ f^{(1)}[X(t)]\epsilon(t) \tag{64}$$

Es más, si el supuesto de Markov (supuesto 6) vale, entonces las probabilidades de transición acotadas para X satisfarán la ecuación (40). Por lo tanto, los procesos con resultados tipo I y tipo II son indistinguibles de los procesos con sólo resultados tipo I.

Trayectorias discontinuas con "eventos raros"

En esta sección concluyendo, el caso general se analiza donde los resultados para $\epsilon(k),\ k = 1, ..., n$ puede ser tipo I, tipo II, o tipo III. Como era verdad para los procesos en la sección 1.2, implícitamente todo los resultados observados de $\epsilon(k)$ serán del tipo I . Sin embargo, a diferencia de los resultados encontrados en la sección 1.2, la posibilidad de eventos raros con resultados tipo III si "importan". Mientras los resultados tipo III con sus probabilidades proporcionales a h son los más "raros" de los resultados admisibles, las dimensiones de estos resultados también son los más grandes. De hecho, puesto que estos resultados son $O(1)$, se sigue que X puede tener cambios no locales en el valor, incluso en un intervalo infinitesimal, y por consiguiente la muestra de trayectoria resultante para X será discontinua. El análisis que demuestra esto y otras propiedades importantes puede ser simplificado omitiendo resultados tipo II y en la representación de las conclusiones alcanzadas en la sección 1.2, esta simplificación puede hacerse sin pérdida de generalidad.

Proposición 4 *Si para $k = 1, ..., n$, por lo menos un posible resultado para $\epsilon(k)$ es un resultado tipo III, entonces la muestra de trayectoria para el precio del valor no será continua.*

Demostración

Sea $Q_k(\delta)$ la probabilidad que $|X(k)-X(k-1)| \geq \delta$ condicional en saber toda la información disponible a partir del tiempo $k-1$. Como en las demostraciones de las proposiciones anteriores, un requisito y condición suficiente para la continuidad de la muestra de trayectoria para X será que, para cada $\delta > 0$, $Q_k(\delta) = o(h)$. Para cada k, supóngase el evento j que denota un resultado tipo III para $\epsilon(k)$ donde $\epsilon(k) = \epsilon_j$ y ϵ_j es $O(1)$. Si p_j es la probabilidad condicional a partir del tiempo $k-1$ que ocurra el evento j , entonces, de (1.5), puede escribirse p_j como $\lambda_j h$ donde $\lambda_j = O(1)$. Se escoje un número θ tal que $0 < \theta < 1$. Se define h^+ por $h^+ = \infty$ si $\alpha_k \epsilon_j \geq 0$ y $h^+ = (\theta - 1)\epsilon_j/\alpha_k$ si $\alpha_k \epsilon_j < 0$. Nótese que $h^+ > 0$, es independiente de h. Se define $\delta^+ \equiv \theta|\epsilon_j|$. Nótese que $\delta^+ > 0$, es independiente de h. Para toda h tal que $0 < h \leq h^+$, se sigue que, si $\epsilon(k) = \epsilon_j$ entonces $|X(k) - X(k-1)| > \delta^+$. Ahora, para $0 < h \leq h^+$ y cualquier δ tal que $0 < \delta \leq \delta^+$, $|X(k) - X(k-1)| > \delta$ si $\epsilon(k) = \epsilon_j$ y por consiguiente $Q_k(\delta) > \lambda_j h = O(h)$ porque $\lambda_j = O(1)$. Aquí, la muestra de trayectoria no es continuo. $\square$

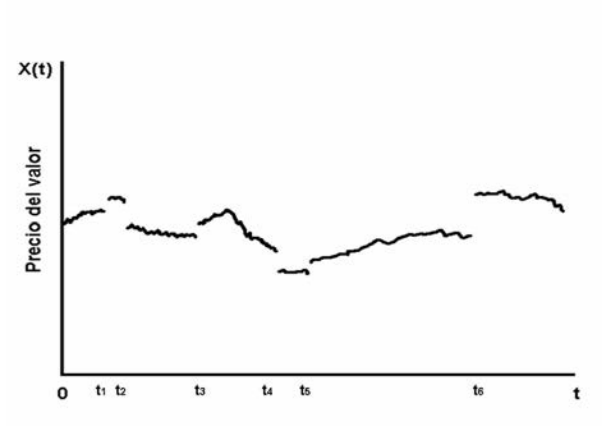

Figura 2: Muestra de trayectorias continuas para el precio de un valor con resultados "Tipo I" y "Tipo III" t_i denota el tiempo al cual los reultados tipo III son hechos

Como se ilustra en la Figura 2, una muestra de trayectoria típica contendrá los cambios principalmente locales o continuos con poca frecuencia cambios no locales o "saltos" que corresponden a resultados tipo III relativamente raros.

Estas discontinuidades fundamentales en la muestra de trayectorias se manifiestan en las propiedades de momento de $X(k) - X(k-1)$. Como los procesos en las secciones 1.1 y 1.2, los primeros y segundos momentos incondicionales de $X(k)-X(k-1)$ son asintóticamente proporcional a h. Sin embargo, a diferencia de los procesos en esas secciones, los Nth momentos absolutos incondicionales,

$2 < N < \infty$, también son asintóticamente proporcionales a h. Es decir,

$$\begin{aligned} I\!E_0\left\{|\epsilon(k)|^N\right\} &= \sum_1^m p_j|\epsilon_j|^N \\ &= O\left(\sum_1^m h^{(N-2)r_j+1}\right) \\ &= O(h) \quad N > 2 \end{aligned} \tag{65}$$

porque $r_j = 0$ para todos los resultados tipo III. Así, todos los momentos absolutos de $X(k) - X(k-1)$ son del mismo orden de magnitud y por consiguiente ninguno de los momentos podrá ser despreciado. Sin embargo, las contribuciones de los resultados tipo I para $\epsilon(k)$ a los momentos superiores que el segundo se mostrarán ser asintóticamente insignificantes. Para mostrar esto junto con los otros resultados, es útil dividir los resultados formalmente para $\epsilon(k)$ en componentes tipo I y tipo III.

Se define la variable aleatoria condicional $u(k)$ por $u(k) \equiv \epsilon(k)/h^{1/2}$ condicional en $\epsilon(k)$ teniendo un resultado tipo I. Igualmente, se define la variable aleatorio condicional $y(k)$ por $y(k) \equiv \epsilon(k)$ condicional en $\epsilon(k)$ teniendo un resultado tipo III. Si $\lambda(k)h$ denota la probabilidad condicional a partir del tiempo $k-1$ que $\epsilon(k)$ tiene un resultado tipo III, entonces, a partir del tiempo $k-1$, se puede escribir a $\epsilon(k)$ como

$$\epsilon(k) = \begin{cases} u(k)h^{1/2} & \text{con probabilidad } 1-\lambda(k)h \\ y(k) & \text{con probabilidad } \lambda(k)h \end{cases} \tag{66}$$

y, por construcción, $u(k),\ y(k)$, y $\lambda(k)$ son todos $O(1)$.

Si $I\!E^y_{k-1}$ denota la esperanza condicional a partir del tiempo $k-1$ en la función de distribución para $y(k)$ y $I\!E^u_{k-1}$ denota la esperanza condicional correspondiente en la distribución para $u(k)$, entonces

$$\overline{y}(k) \equiv I\!E^y_{k-1}\left\{y(k)\right\} = I\!E_{k-1}\left\{\epsilon(k)|\text{resultado tipo III}\right\}$$

y

$$\overline{u}(k)h^{1/2} \equiv h^{1/2} I\!E^u_{k-1}\left\{u(k)\right\} = I\!E_{k-1}\left\{\epsilon(k)|\text{resultado tipo I}\right\}$$

Puesto que $I\!E_{k-1}\left\{\epsilon(k)\right\} = 0$, se sigue inmediatamente de las propiedades de esperanza condicional que

$$\begin{aligned} \overline{u}(k) &= -\frac{\lambda(k)\overline{y}(k)h^{1/2}}{1-\lambda(k)h} \\ &= -\lambda\overline{y}h^{1/2} + O(h) \end{aligned} \tag{67}$$

donde se suprime la dependencia explícita de $u,\ y$, y λ en k siempre que no hay ninguna ambigüedad acerca del tiempo.

Si σ_k^2 es la varianza condicional por unidad de tiempo para $\epsilon(k)$, entonces $\mathbb{E}_{k-1}\left\{\epsilon^2(k)\right\} = \sigma_k^2 h$, y se sigue que

$$\begin{aligned}\sigma_u^2 &= \frac{\sigma_k^2 - \lambda\sigma_y^2}{1-\lambda h} \\ &= \sigma_k^2 - \lambda\sigma_y^2 + O(h) \qquad (68)\end{aligned}$$

donde σ_u^2 es la varianza condicional de $u(k)$ y σ_y^2 es la varianza condicional de $y(k)$. Nótese que σ_k^2, σ_u^2, y σ_y^2 son todas $O(1)$. Aún más, para $N > 2$, se sigue que

$$\mathbb{E}_{k-1}\left\{\epsilon^N(k)\right\} = \lambda(k)\mathbb{E}_{k-1}^y\left\{y^N(k)\right\}h + o(h) \qquad (69)$$

Así, mientras ambos resultados tipo I y tipo III contribuyen significativamente a la media y varianza de $\epsilon(k)$, las contribuciones de los resultados tipo I a los momentos superiores de $\epsilon(k)$ son asintóticamente insignificante.

Como se hizo en las secciones anteriores, se completará el análisis examinando las características de distribución de variables aleatorias que son funciones del precio del valor. Como en la sección 1.1, sea $F(t)$ una variable aleatoria dada por la regla que $F(t) = f(X,t)$ si $X(t) = X$ donde f es una función C^2 con derivadas parciales de orden tres acotadas.

Para un resultado dado de $y(k) = y$, se tiene por el teorema de Taylor que

$$\begin{aligned}f\left[X(k-1)+\alpha_k h + y, k\right] &= f\left[X(k-1)+y, k-1\right] \\ &\quad + f_1\left[X(k-1)+y,k-1\right]\alpha_k h \\ &\quad + f_2\left[X(k-1)+y,k-1\right]h + o(h)\end{aligned} \qquad (70)$$

donde, como en la sección 1.1, los subíndices en f denotan las derivadas parciales.

Similarmente, para resultados dados de $u(k) = u$, se tiene que

$$\begin{aligned}f\left[X(k-1)+\alpha_k h + uh^{1/2}, k\right] &= f\left[X(k-1), k-1\right] \\ &\quad + f_1\left[X(k-1),k-1\right]\left(\alpha_k h + uh^{1/2}\right) \\ &\quad + f_2\left[X(k-1),k-1\right]h \\ &\quad + \tfrac{1}{2}f_{11}\left[X(k-1),k\right]u^2 h + o(h) \qquad (71)\end{aligned}$$

Por las propiedades de esperanza condicional, se tiene que

$$\begin{aligned}\mathbb{E}_{k-1}\left\{F(k)-F(k-1)\right\} &= \lambda(k)h\mathbb{E}_{k-1}^y\left\{F(k)-F(k-1)\right\} \\ &\quad + [1-\lambda(k)h]\mathbb{E}_{k-1}^u\left\{F(k)-F(k-1)\right\}\end{aligned} \qquad (72)$$

Sustituyendo de (70) y (71) en (72) y eliminando las representaciónes explícitas de términos que son $o(h)$, se puede volver a escribir (72) como

$$\begin{aligned}
I\!E_{k-1}\left\{F(k)-F(k-1)\right\} \;=\; & (\tfrac{1}{2}f_{11}\left[X(k-1),k-1\right]\sigma_u^2 \\
& +f_1\left[X(k-1),k-1\right](\alpha_k-\lambda\overline{y}) \\
& +f_2\left[X(k-1),k-1\right] \\
& +\lambda I\!E_{k-1}^{y}\left\{f\left[X(k-1)+y(k),k-1\right]\right. \\
& \left.-\; f\left[X(k-1),k-1\right]\right\})h+o(h) \qquad (73)
\end{aligned}$$

Si en una forma correspondiente a la ecuación (22) se define la esperanza condicional del cambio en F por unidad de tiempo como

$$\mu_k \equiv I\!E_{k-1}\left\{F(k)-F(k-1)|\right\}/h$$

entonces dividiendo (73) por h y tomando el límite cuando $h\rightarrow 0$, se puede escribir la esperanza condicional instantánea del cambio en F por unidad de tiempo, $\mu(t)$, como

$$\begin{aligned}
\mu(t) \;=\; & \tfrac{1}{2}f_{11}\left[X(t),t\right]\sigma_u^2+f_1[X(t),t][\alpha(t)-\lambda(t)\overline{y}(t)]+f_2[X(t),t] \\
& +\lambda(t)I\!E_t^{y}\left\{f[X(t)+y(t),t]-f[X(t),t]\right\} \qquad (74)
\end{aligned}$$

Nótese que en el caso especial cuando $\lambda(t)=0$ y no hay ningún resultado tipo III, la expresión para $\mu(t)$ en (74) se reduce a la forma correspondiente de (22) deducida en la sección 1.1.

En una forma similar, los momentos condicionales superiores para el cambio en F pueden escribirse como

$$\begin{aligned}
& I\!E_{k-1}\left\{[F(k)-F(k-1)]^2\right\} = \\
& \qquad (\lambda I\!E_{k-1}^{y}\left\{f[X(k-1)+y(k),k-1]-f[X(k-1),k-1]\right\}^2 \\
& \qquad +\; f_1^2[X(k-1),k-1]\sigma_u^2)h+o(h) \qquad (75)
\end{aligned}$$

y, para $N>2$,

$$\begin{aligned}
& I\!E_{k-1}\left\{[F(k)-F(k-1)]^N\right\} = \\
& \qquad \lambda I\!E_{k-1}^{y}\left\{f[X(k-1)+y(k),k-1]-f[X(k-1),k-1]\right\}^N h+o(h) \qquad (76)
\end{aligned}$$

Como era el caso para los momentos de $X(k)-X(k-1)$, todos los momentos de $F(k)-F(k-1)$ son del mismo orden de magnitud, y sólo los resultados tipo III contribuyen significativamente a los momentos superiores que el segundo. Ahora, las únicas características de $u(k)$ que importan son sus primeros dos momentos.

Si ahora se reintroduce el supuesto 6 que el proceso estocástico para $X(t)$ es Markov, entonces $\lambda(k)=\lambda[X(k-1),k-1]$; $\alpha_k=\alpha_k[X(k-1),k-1]$; $\sigma_u^2(k)=$

$\sigma_u^2[X(k-1), k-1]$; y $g_k(y)$, la función de densidad condicional para $y(k)$, puede escribirse como $g[y(k); X(k-1), k-1]$. Como se definió en (39) de la sección 1.1, sea $p(x,t)$ que denota la densidad de probabilidad condicional para $X(T) = X$ al momento T, condicional en $X(t) = X$. Para X y T fijas, $p[X(t), t]$ es una variable aleatoria que está en función del precio del valor al momento t. Ahora, la esperanza instantánea del cambio en p por unidad de tiempo satisfará (74). Sin embargo, puesto que p es una densidad de probabilidad, su cambio esperado es cero. Sustituyendo la condición que $\mu(t) = 0$ en (74), se tiene que p debe satisfacer

$$\begin{aligned} 0 &= \tfrac{1}{2}\sigma_u^2 p_{11}(x,t) + (\alpha - \lambda\overline{y})p_1(x,t) + p_2(x,t) \\ &\quad +\lambda \int p(x+y,t)g(y;x,t)dy \end{aligned} \tag{77}$$

que es una ecuación diferencial parcial lineal diferente para las probabilidades de transición de $p(x,t)$. Del, conocimiento de las funciones σ_u^2, α, λ, y g es suficiente determinar la distribución de probabilidad para el cambio en X entre cualquiera dos fechas.

Por lo tanto, sin pérdida de generalidad, se podrá describir siempre la dinámica de $X(t)$ por la ecuación diferencial estocástica

$$dX(t) = \{\alpha[X(t),t] - \lambda[X(t),t]\overline{y}(t)\}\, dt \; + \; \sigma[X(t),t]dZ(t) \; + \; y(t)dq(t) \tag{78}$$

donde α es la esperanza instantánea del cambio en X por unidad de tiempo; σ^2 es la varianza instantánea del cambio en X, condicional en que el cambio sea un resultado tipo I; λ es la probabilidad por unidad de tiempo que el cambio en X es un resultado tipo III; y $y(t)$ es la variable aleatoria para el cambio en X, condicional en tener de cambio un resultado tipo III y $dq(t)$[7]. La representación de la diferencial estocástica en (78) realmente se define por una integral estocástica $X(T) - X(0) = \int_0^T dX(t)$. Claro, si $\lambda \equiv 0$ y sólo resultados tipo I pueden ocurrir, entonces (78) se reduce a (45) en la sección 1.1.

En una forma similar, puede mostrarse que la representación de la diferencial estocástica para F puede escribirse como

$$\begin{aligned} dF(t) &= \left\{\tfrac{1}{2}\sigma^2 f_{11}[X(t),t] + (\alpha - \lambda\overline{y})f_1[X(t),t] + f_2[X(t),t]\right\} dt \\ &\quad +\sigma f_1[X(t),t]dZ(t) + \{f[X(t)+y(t),t] - f[X(t),t]\}\, dq(t) \end{aligned} \tag{79}$$

Así, si la dinámica de $X(t)$ puede describirse por una superposición de los procesos de difusión y Poisson, entonces la dinámica de funciones bien portadas de $X(t)$ pueden describirse de la misma forma. Así, (79) proporciona la regla de la transformación que corresponde al lema de Itô para los procesos de difusión.

En resumen, si la estructura económica a ser analizada es tal que los supuestos 1-5 valen, entonces la dinámica del precio del valor en los modelos siempre

[7]Es la diferencial de un proceso Poisson, que es definido en el capítulo siguiente.

pueden describirse por una “mezcla” de procesos de difusión con muestra de trayectorias continuas y procesos Poisson sin pérdida de generalidad. La componente del proceso de difusión describe los cambios locales frecuentes en los precios. La componente del proceso Poisson se usa para capturar esos eventos raros cuando las variables tienen los cambios no locales y el precio del valor “salta”.

Teoría del precio justo

En el presente capítulo se pretende deducir un conjunto de restricciones para obtener una forma de valuación de opciones, esas restricciones son condiciones necesarias para que sean consistentes con la teoría del precio justo. También se introducen supuestos adicionales para examinar y extender la teoría del precio justo elaborada por Black-Scholes. Estas restricciones y supuestos están basadas en que las preferencias de los inversionistas tiende hacia una mejor posición siempre.

Restricciones y supuestos en las opciones

Una opción call tipo "Americana" es un valor, emitido por un individuo, dando a su dueño el derecho a comprar una porción de acción a un precio dado en o antes de una fecha dada. Una opción put tipo "Americana" da a su dueño el derecho de vender una porción de acción a un precio dado en o antes de una fecha dada. Una opción tipo "Europea" tiene los mismos términos como su contraparte "Americana" excepto que no puede ser ejercida antes del último día del contrato.

La notación usada será la siguiente: $G(X,t;E)$ es el valor de una opción put Americana con precio de ejercicio E y t años antes de expirar, donde el precio por porción de acción común es X; $g(X,t;E)$ es el valor de su contraparte Europea.

Usando la definición anterior, se tiene que, a la fecha de vencimiento,

$$G(X,0;E) = g(X,0;E) = \text{máx}(0, E - X)$$

Supuesto 1 *Una condición necesaria para la Teoría del precio justo de una opción es que la opción se valué tal que no domine* [1] *ni lo domine un valor.*

Para determinar el precio justo de una opción put Europea, las posiciones de dos portafolios son examinadas. Considerando una posición larga en la acción común a X pesos, una posición larga en un put Europeo a t años a $g(X,t;E)$ pesos, y un préstamo de $EP'(t)$ pesos donde $P'(t)$ es el valor actual de un peso pagadero t años a la tasa de prestamo. El valor del portafolio t años desde ahora con el precio de acción a X^* será $X^* + (E - X^*) - E = 0$ si $X^* \geq E$, y $X^* + 0 - E = X^* - E$ si $X^* > E$. La estructura de la ganacia neta es idéntica en cada estado para una opción call Europea con el mismo precio de ejercicio y fecha de vencimiento. Ahora, para evitar que la opción call sea un valor dominado, la opción put y call deben ser preciados de manera que

$$g(X,t;E) + X - EP'(t) \geq f(X,t;E)$$

[1] Un valor A es *dominante* sobre un valor B si, en alguna fecha conocida en el futuro, el rendimiento en A excederá al rendimiento en B para algún posible estado del mundo, y será por lo menos tan grande como en B en todos los posibles estados del mundo.

Por otro lado, una posición larga en una call Europea a t años, una posición corta en una acción común a un precio X, y prestando $EP(t)$ pesos. El valor del portafolio a t años desde ahora con precio de acción a X^* será $0 - X^* + E = E - X^*$ si $X^* \geq E$, y $(X^* - E) - X^* + E = 0$ si $X^* > E$. La estructura de la ganancia neta es idéntica en cada estado para una opción put Europea con el mismo precio de ejercicio y fecha de vencimiento. Si la opción put no es un valor dominado, entonces

$$f(X, t; E) - X + EP(t) \leq g(X, t; E)$$

Teorema 1 *Si el supuesto 1 vale y si la tasa de prestamo y de pedir prestado son iguales (esdecir, $P(t) = P'(t)$), entonces*

$$g(X, t; E) = f(X, t; E) - X + EP'(t)$$

Así, el valor de una opción put Europea evaluado justamente es determinado una vez que se tiene una teoria racional del valor de una opción call. La forma deducida en el teorema anterior es idéntica con la ecuación de Black-Scholes cuando la tasa libre de riesgo r es constante (es decir, $P(t) = \exp(-rt)$).

Supuesto 2 *No hay costos de transacción o impuestos. El intercambio toma lugar continuamente y el pedir préstamo y las ventas-cortas[2] son permitidas sin restricción [3].*

Supuesto 3 *Los valores y mercados de intercambio son "suficientemente perfectos", y la dinámica de los rendimientos de los recursos son tal que los valores son valuadas para satisfacer el Capital Asset Pricing Model.*

Supuesto 4 *Existe un mercado de intercambio donde algún inversionista e institución pueden pedir prestado o prestar a la misma tasa de interés r, la cual es constante a través del tiempo. Por consiguiente, para evitar arbitraje por estos inversionistas, $r \geq R + s$.*

Supuesto 5 *A pesar de que los depósitos individuales son del tipo demanda, se asume que las dinámicas para los depósitos agregados para una institición dada, D, son no estocásticos y son descritos por $dD = gDdt$ donde el porcentaje de crecimiento en los depósitos, g, es una constante conocida.*

---

[2] Se denomina **hacer corto** o **ventas-cortas** cuando se puede vender un contrato en el mercado, *incluso si no se le posee*. Un inversionista actúa como emisor de un contrato y tiene las mismas obligaciones del emisor. Por ejemplo, la actividad de vender un valor básico tal como una acción sin *poseerla* en *primer lugar* y luego comprarla para entregarla.

[3] Los supuestos de préstamo y ventas-cortas sin restricción pueden ser debilitados sin cambio en los resultados obtenidos por la división del portafolio creado en dos portafolios: uno conteniendo la acción ordinaria y el otro conteniendo el warrant más una posición larga en el depósito. Entonces si se acepta el supuesto 1, las ecuaciones de la sección actual son inmediatas. Véase [**?**]

Modelo Black-Scholes

La idea principal del modelo de Black-Scholes fue expuesta en la última parte de la sección 1.1, en la cual se llego a la conclusión de que la dinámica del precio de la acción se compone de un movimiento Browniano geométrico. En otras palabras, los movimientos del mercado deben obedecer la ecuación diferencial estocástica por el movimiento Browniano geométrico. Además aplicando el Lema de Itô se deduce la siguiente ecuacón

$$df = \left(\frac{\partial f}{\partial X} \alpha X + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial X^2} \sigma^2 X^2 \right) dt + \frac{\partial f}{\partial X} \sigma X dZ(t) \tag{1}$$

Construcción de un portafolio libre de riesgo

El portafolio se constituirá de una cantidad de opciones y una cantidad del bien subyacente, para tener un portafolio apropiado. El valor del portafolio es

$$\prod = f - kX \tag{2}$$

donde k es el valor que representa el número de unidades de acciones vendidas por cada unidad de opción. El cambio $d\prod$ en el valor del portafolio en un intervalo de tiempo esta dado por

$$d\prod = df - kdX$$

Sustituyéndose las ecuaciones (1) y (**??**) en la ecuación anterior, se tiene que

$$\begin{aligned} d\prod &= \left[\left(\tfrac{\partial f}{\partial X} \alpha X + \tfrac{\partial f}{\partial t} + \tfrac{1}{2} \tfrac{\partial^2 f}{\partial X^2} \sigma^2 X^2 \right) dt + \tfrac{\partial f}{\partial X} \sigma X dZ(t) \right] - k \left(\alpha X dt + \sigma X dZ(t) \right) \\ &= \left[\left(\tfrac{\partial f}{\partial X} \alpha X + \tfrac{\partial f}{\partial t} + \tfrac{1}{2} \tfrac{\partial^2 f}{\partial X^2} \sigma^2 X^2 - k\alpha X \right) dt \right] + \left(\tfrac{\partial f}{\partial X} \sigma X - k\sigma X \right) dZ(t) \end{aligned} \tag{3}$$

Para obtener un portafolio libre de riesgo es necesario eliminar la componente aleatoria $dZ(t)$ por consiguiente

$$k = \frac{\partial f}{\partial X} \tag{4}$$

Al sustituirse k la razón de cobertura en (2) se tiene

$$\begin{aligned} d\prod &= \left[\left(\tfrac{\partial f}{\partial X} \alpha X + \tfrac{\partial f}{\partial t} + \tfrac{1}{2} \tfrac{\partial^2 f}{\partial X^2} \sigma^2 X^2 - \tfrac{\partial f}{\partial X} \alpha X \right) dt \right] + \left(\tfrac{\partial f}{\partial X} \sigma X - \tfrac{\partial f}{\partial X} \sigma X \right) dZ(t) \\ &= \left(\tfrac{\partial f}{\partial t} + \tfrac{1}{2} \tfrac{\partial^2 f}{\partial X^2} \sigma^2 X^2 \right) dt \end{aligned} \tag{5}$$

El supuesto de que el portafolio debe ser libre de riesgo implica que debe ganar instantáneamente la misma tasa de rendimiento como otro valor a corto plazo libre de riesgo. Es decir

$$d\prod = r\prod dt$$

donde r es la tasa de rendimiento libre de riesgo. Sustituyéndose (2), (4) y (5) se obtiene

$$\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}\sigma^2 X^2\right) dt = r(f - \frac{\partial f}{\partial X}X)dt$$

$$\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}\sigma^2 X^2 = rf - r\frac{\partial f}{\partial X}X$$

Agrupándose términos se tiene la ecuación diferencial estocástica de Black and Scholes

$$\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}\sigma^2 X^2 + r\frac{\partial f}{\partial X}X - rf = 0 \tag{6}$$

Proceso Poisson

El así llamado cálulo estocástico del proceso Poisson esta basado en la noción de integral con respecto a un proceso de conteo.

Esta sección introduce *procesos de conteo (o eventos aleatorios)* de los cuales el ejemplo más simple es el *proceso Poisson homogéneo*. Un proceso de eventos aleatorios es, rigurosamente hablando, un conjunto aleatorio contable de puntos de la recta real. En general, los procesos de conteo aparecen en los modelos estocásticos donde el estado de un sistema esta cambiando por la ocurrencia de algunos eventos.

Definición 1 *Un proceso de conteo en* $\mathbb{R}_+$ *es una sucesión* $\{T_n\}_{n\geq 0}$ *de variables aleatorias no negativas tal que, casi seguramente: (i)* $T_0 \equiv 0$, *(ii)* $0 < T_1 < T_2 < \cdots$, *y (iii)* $\lim_{n\uparrow\infty} T_n = \infty$

Por otro lado, el Proceso Poisson surge naturalmente modelando el número de ocurrencias de eventos aleatorios independientes en un intervalo de tiempo de longitud h (donde $h = (t, t + h)$ tan pequeño como se quiera), en donde en cada caso el número de ocurrencias en un intervalo es en promedio, el producto de h con un número fijo λ la cual es interpretada como la freuencia promedio de ocurrencias para cada evento. Así en un intervalo pequeño se espera encontrar a lo más una ocurrencia, la probabilidad de exactamente una ocurrencia en un intervalo pequeño de longuitud h es aproximadamente λh. Si admás se asume que las ocurrencias no están relacionadas, se tiene que esto es conocido como un Proceso Poisson.

Definición 2 *Sea $N(t)$ el número de ocurrencias de un evento aleatorio en un intervalo de tiempo $(0,t]$. Si $N(t)$ satisface los siguientes supuestos, debe llamarse a $N(t)$ un Proceso Poisson de parámetro λ*

i) Existe $\lambda > 0$ tal que para cualquier $t \geq 0$ y $h > 0$

- **a)** $\mathbb{P}\left[N(t+h) - N(t) = 0\right] = 1 - \lambda h + o(h)$
- **b)** $\mathbb{P}\left[N(t+h) - N(t) = 1\right] = \lambda h + o(h)$
- **c)** $\mathbb{P}\left[N(t+h) - N(t) = k\right] = o(h), \quad k \geq 2$

ii) Si N_h es el número de ocurrencias en el intervalo de tiempo $h = (t, t+h]$, entonces $N_{h_1}, N_{h_2}, \ldots, N_{h_k}$ son independientes siempre que $h_1, h_2, \ldots, h_k$ sean disjuntos.

La condición *i)* es la propiedad de incrementos independientes del proceso Poisson. Esto implica en particular que para algún intervalo $(a,b]$, la variable aleatoria $N(a,b]$ es independiente de $N(s]$, $s \in (0,a]$. Por esta razón, al proceso Poisson algunas veces es llamado procesos sin memoria. Los incrementos de un proceso Poisson homogéneo no tienen memoria del pasado.

Para encontrar la distribución de probabilidad de $N(t)$, se introduce la notación

$$\mathbb{P}_k(t) = \mathbb{P}\left[N(t) = k\right] \tag{7}$$

así que por los supuestos i) y ii),

$$\begin{aligned} \mathbb{P}_0(t+h) &= \mathbb{P}\left[N(t+h) = 0\right] \\ &= \mathbb{P}\left[N(t) = 0, N(t+h) = 0\right] \\ &= \mathbb{P}\left[N(t) = 0\right] \mathbb{P}\left[N(t+h) = 0\right] \\ &= \mathbb{P}_0(t)\left[1 - \lambda h + o(h)\right] \end{aligned}$$

Por consiguiente

$$\frac{\mathbb{P}_0(t+h) - \mathbb{P}_0(t)}{h} = -\lambda \mathbb{P}_0(t) + \frac{o(h)}{h}\mathbb{P}_0(t)$$

el cual, cuando $h \to 0$, produce

$$\frac{d\mathbb{P}_0(t)}{dt} = -\lambda \mathbb{P}_0(t)$$

La solución a esta ecuación diferencial, notando que $\mathbb{P}_0(0) = 1$, es

$$\mathbb{P}_0(0) = e^{-\lambda t}$$

Obteniéndose la fórmula general

$$\mathbb{P}_k(t) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

Por lo tanto, $N(t) \sim \text{Poisson}(\lambda t)$.

Integral de Poisson

Sea $\{T_n\}_{n\geq 1}$ un proceso punto en $(0,\infty]$ y $\{N(t)\}_{t\geq 0}$ la función de conteo asociada. Sea $(a,b]\subset\mathbb{R}_+$ un intervalo finito arbitrario y $\{Z(t)\}_{t\geq 0}$ un proceso estocástico arbitrario con valores complejos.

Definición 3 *La siguiente ecuación es llamada la integral de Stieltjes-Lebesgue, o la integral de conteo, de Z con respecto a N en $(a,b]$. Claramente, esta definición podría darse con algún conjunto acotado $C\subset\mathbb{R}_+$ reemplazando a $(a,b]$.*

$$\int_{(a,b]} Z(t)dN(t) \overset{def}{=} \sum_{n\geq 1} Z(T_n)\mathbf{1}_{(a,b]}(T_n) \tag{8}$$

Para intervalos acotados, la definición de la integral no presenta dificultades puesto que la suma involucrada tiene un número finito de términos.

Nótese que si $Z(t)$ esta definida en $(a,b]$ por $Z(t)=c_j\in\mathbb{C}$ en $(t_j,t_{j+1}]$ para toda $k\in[0,K]$, dode $a=t_0<t_1<\cdots<t_{K+1}=b$, entoces

$$\int_{(a,b]} Z(t)dN(t) = \sum_{j=0}^{K} c_j[N(t_{j+1})-N(t_j)]$$

Teorema 2 *Para $j\in I$ fija, si $\{Z(t)\}_{t\geq 0}$ es un proceso estocástico complejo con trayectorias continuas por la izquierda, y tal que para toda $t\geq 0$, la variable aleatoria $Z(t)$ es una fucncional de $\underline{Y}$, de $\{N_k\}_{k\in I,k\neq j}$, y de la restricción de N_j a $(0,t]$. Supóngase que al menos una de las siguientes condiciones se cumple:*

i) *$Z(t)_{t\geq 0}$ es no negativa.*

ii) *$Z(t)_{t\geq 0}$ toma valores complejos, y $I\!E[\int_0^\infty |Z(t)|N_j(dt)]$ o $I\!E[\int_0^\infty |Z(t)|\lambda_j dt]$ es finita.*

Entonces la fórmula suavizada es:

$$I\!E\left[\int_0^\infty Z(s)dN_j(s)\right] = I\!E\left[\int_0^\infty Z(s)\lambda_j ds\right] \tag{9}$$

Demostración

La demostración puede consultarse en [**?**] $\square$

Modelo Black-Scholes con discontinuidades

El cambio total en el precio de una acción es situado a ser la composición de dos tipos de cambios:

1. El primero es la variación "*normal*" en el precio debido, por ejemplo, a una temporada inestable entre la demanda y la oferta, cambios en las tasas de capitalización, cambios en las perspectivas económicas, u otra información nueva que cause cambios marginales en el valor de la acción. En esencia, el impacto de tal información por unidad de tiempo en el precio de la acción esta produciendo un cambio marginal en el precio. Esta componente es modelada por un movimiento Browniano geométrico con una constante de variación por unidad de tiempo y tiene una muestra de trayectorías continuas.

2. El segundo es la variación "*anormal*" en el precio debido a la llegada de información importante nueva acerca de la acción que tiene un efecto más marginal en el precio. Esta componente es modelada por un proceso de salto reflejando el impacto no marginal de la información.

Así como una vez las dinámicas estan postuladas a ser un proceso continuo, el proceso natural para la componente continua del cambio del precio de la acción es un proceso de Wiener, así el proceso para la componente de salto es un proceso Poisson.

Dinámica del precio de una acción y una opción

Dado que el evento Poisson ocurre (es decir, arriba información importante sobre la acción), entonces se genera un "bosquejo" de una distribución para determinar el impacto de esta información en el precio de la acción. Esto es, si $X(t)$ es el precio de la acción al tiempo t y Y es la variable aleatoria que describe este bosquejo, entonces, omitiendo la parte continua, el precio de la acción al tiempo $t+h$, $X(t+h)$, será la variable aleatoria $X(t+h) = X(t)Y$, dado que ocurrió un arribo de información entre t y $t+h$. Se asumirá a lo largo que Y tiene una medida de probabilidad con soporte compacto y $Y \geq 0$. Además, los $\{Y\}$ de sucesivos bosquejos son independientes e idénticamente distribuidos.

La composición de los rendimientos del precio de la acción pueden ser escritos como una ecuación diferencial estocástica, condicionada en $X(t) = X$, como

$$dX = (\alpha - \lambda k)X dt + \sigma X dZ + X dq \tag{10}$$

donde α es el rendimiento esperado instantáneo en la acción; σ^2 es la variación instantánea del rendimiento, condicionada en ningún arribo de información importante nueva (i.e. el evento Poisson no ocurre); dZ es un proceso Wiener; $q(t)$ es el proceso Poisson independiente descrito previamente; $dq(t)$ y dZ se asumen independientes; λ es el número medio de arrivos por unidad de tiempo; $k = \{Y - 1\}$ donde $Y - 1$ es el porcentaje de cambio de la variable aleatoria en el precio de la acción si el evento Poisson ocurre.

La parte σdZ describe la parte instantánea del rendimiento inesperado debido a las oscilaciones "normales" del precio, y la parte dq describe la parte debida a las oscilaciones "anormales" del precio.

Habiéndose establecido la dinámica del precio de una acción, se analizará la dinámica del precio de una opción. Supóngase que el precio de una opción W se puede escribir como una función diferencial del precio de una acción y del tiempo: es decir, $W(t) = F(X, t)$. Si el precio de la acción sigue la dinámica descrita, entonces la dinámica del rendimiento de la opción puede escribirse en una forma similar como

$$dW = (\alpha_W - \lambda k_W)W dt + \sigma_W W dZ + W dq_W \tag{11}$$

donde α_W es el rendimiento esperado instantáneo en la opción; σ_W^2 es la variación instantánea del rendimiento, condicionada en que no ocurra el evento Poisson. $q_W(t)$ es el proceso Poisson independiente con parámetro λ. $k_W = \mathbb{E}\{Y_W - 1\}$ donde $Y_W - 1$ es el porcentaje de cambio de la variable aleatoria en el precio de la opción si el evento Poisson ocurre.

Usando el lema de Itô para la parte continua y discontinua, se tienen las siguientes relaciones

$$\alpha_W \equiv \frac{\frac{1}{2}\sigma^2 X^2 F_{11}(X,t) + (\alpha - \lambda k) X F_1(X,t) + F_2(X,t) + \lambda \mathbb{E}\{F(XY,t) - F(X,t)\}}{F(X,t)} \tag{12}$$

$$\sigma_W \equiv \frac{F_1(X,t)\sigma X}{F(X,t)} \tag{13}$$

Además, el proceso Poisson para el precio de la opción, $q_W(t)$, es dependiente funcionalmente en el proceso Poisson para el precio de la acción, $q(t)$. Es decir, el evento Poisson para el precio de la opción ocurre si y sólo si el evento Poisson para el precio de la acción ocurre. Más aun, si el evento Poisson para la acción ocurre y la variable aleatoria Y toma el valor $Y = y$, entonces el evento Poisson para la opción ocurre y la variable aleatoria Y_W toma el valor $F(Xy, t)/F(X, t)$, i.e. $Y_W \equiv F(XY, t)/F(X, t)$.

Construcción de un portafolio

Considérese un portafolio el cual posee la acción, la opción y el bien sin riesgo (con rendimiento r por unidad de tiempo) en proporciones w_1, w_2 y w_3 donde $\sum_{j=1}^{3} w_j = 1$. Si P es el valor del portafolio, entonces la dinámica del rendimiento en el portafolio puede ser escrita como

$$dP = (\alpha_p - \lambda k_p)P dt + \sigma_p P dZ + P dq_p \tag{14}$$

donde α_p es el rendimiento esperado instantáneo en el portafolio; σ_p^2 es la variación instantánea del rendimiento, condicionada en que no ocurra el evento Poisson. $q_p(t)$ es el proceso Poisson independiente con parámetro λ. $k_p = \mathbb{E}\{Y_p - 1\}$ donde $Y_p - 1$ es el porcentaje de cambio de la variable aleatoria en el valor del portafolio si el evento Poisson ocurre.

De (10) y (11), se tiene que

$$\begin{aligned} \alpha_p &\equiv w_1(\alpha - r) + w_2(\alpha_W - r) + r && (15) \\ \sigma_p &\equiv w_1\sigma + w_2\sigma_W && (16) \\ Y_p - 1 &\equiv w_1(Y-1) + w_2(Y_W - 1) && (17) \end{aligned}$$

donde $w_3 = 1 - w_1 - w_2$.

Inspeccionando la ecuación (17) muestra que no existe un conjunto con porciones (w_1, w_2)[4] de un portafolio tal que elimine el riesgo del "salto" (i.e. hacer a $Y_p \equiv 1$). La razón es que el portafolio mixto es una operación *lineal* y el precio de la opción es una función *no lineal* del precio de la acción. Por consiguiente, en el caso cuando $\sigma^2 = 0$ y Y sea un proceso Poisson, entonces para cualquier w_1 y w_2, $Y_p - 1$ no asumirá valores cero para algún posible valor de Y.

Sin embargo, se pueden analizar las características del rendimiento en un portafolio con cobertura. Sea bajo las condiciones anteriores, P^* el valor del portafolio con "cobertura". Entonces, de (14) se tiene que

$$dP^* = (\alpha_p^* - \lambda k_p^*)P^* dt + P^* dq_p^* \tag{18}$$

Notése que el rendimieto del portafolio es un proceso de salto "puro" puesto que la parte continua de los movimientos del precio de la acción y la opción han sido"cubiertos". Más aun, la ecuación (18) puede ser escrita como

$$\begin{aligned} dP^* = &\quad (\alpha_p^* - \lambda k_p^*)P^* dt && \text{si no ocurre el evento Poisson} \\ dP^* = &\quad (\alpha_p^* - \lambda k_p^*)P^* dt + P^*(Y_p^* - 1) && \text{si ocurre el evento Poisson} \end{aligned} \tag{19}$$

De (19) es fácil ver que, "la mayoría de las veces", el rendimiento en el portafolio será predecible y su rendimiento será $\alpha_p^* - \lambda k_p^*$. Sin embargo, en promedio, una vez cada $1/\lambda$ unidades de tiempo el valor del portafolio tomará un inesperado salto. Además, se puede observar adicionalmente las características cualitativas del rendimiento. Es decir, de (17) y (16),

$$Y_p^* - 1 = w_2^* \left[(Y_W - 1) - \frac{\sigma_W}{\sigma}(Y - 1) \right] \tag{20}$$

[4]En el análisis del modelo Black-Scholes donde $\lambda = 0$. El rendimiento del portafolio puede ser hecho menos riesgoso escogiendo $w_1 = w_1^*$ y $w_2 = w_2^*$ tal que $w_1^*\sigma + w_2^*\sigma_W = 0$. Este hecho, debe ser para evitar el arbitraje del rendimiento esperado en el portafolio con porcentajes w_1^* y w_2^* iguales a la tasa r. De (12), (13) y (21) se puede obtener la ecuación diferencial de Black-Schole.

Por la estricta convexidad del precio de la opción en el precio de la acción, $(Y_W - 1) - (\sigma_W/\sigma)(Y - 1)$ es positivo para cualquier valor de Y. Ahora, si w_2^* es positivo, entonces $Y_p^* - 1$ será positivo, y el rendimiento inesperado en el portafolio con cobertura será siempre positivo. Si $w_2^* < 0$, entonces el rendimiento inesperado será negativo. Más aun, el signo de k_p^* será el mismo que el signo de w_2^*.

Si el origen de los saltos es cualquier información, entonces la componente de salto del rendimiento de la acción representará un riesgo "no sistemático", es decir, la componente de salto no se correlacionará con el mercado. Supóngase que esto es generalmente cierto para las acciones. Entonces es cierto para el rendimiento del portafolio con cobertura P^*. La dinámica del rendimiento en la ecuación (18) muestra que sólo el origen de la incertidumbre en el rendimiento es la componente de salto del la acción. Pero por hipótesis, tal componente representa sólo riesgo no sistemático, y por consiguiente el "coeficioente de susceptibilidad" ("β''") es cero, entonces el rendimiento esperado debe ser igual a la tasa libre de riesgo. Por lo tanto, $\alpha_p^* = r$, pero, de (15) implica que $w_1^*(\alpha - r) + w_2^*(\alpha_W - r) = 0$, se tiene que

$$\frac{\alpha - r}{\sigma} = \frac{\alpha_W - r}{\sigma_W} \tag{21}$$

Por lo tanto, de (21), (12) y (13) implica que F debe satisfacer

$$0 = \frac{1}{2}\sigma^2 X^2 F_{11}(X,t) + (r - \lambda k) X F_1(X,t) + F_2(X,t) - rF + \lambda\epsilon\left\{F(XY,t) - F(X,t)\right\} \tag{22}$$

sujeto a las siguientes condiciones

$$\begin{aligned} F(0,\tau) &= 0 && (23)\\ F(X,0) &= \max(0, X - E) && (24) \end{aligned}$$

donde E es el precio de ejerccio de la opción.

Notése que (22) no depende de α, en lugar de ello, como en el caso del Black-Scholes estándar, sólo aparece la tasa de interés r. Más aún, (22) se reduce a la ecuación (6) si $\lambda = 0$ (es decir, si no hay saltos).

Por lo tanto, la teoría desarrollada por Black y Scholes (1973) es particularmente atractiva por que es una formulación de equilibrio completa del problema del precio justo y la fórmula final es una función de variables "observables", haciendo al modelo sujeto a las pruebas empíricas directas. Se puede observar que Black y Scholes asumen que:

1. La forma estándar de Sharpe-Linter-Mossin *Capital Asset Pricing Model* (CAPM) es valida para intercambios intertemporales, y que los intercambios toman lugar continuamente en el tiempo;

2. La tasa de interés del mercado r es conocida y fija a través del tiempo;

3. No hay dividendos o cambios en el precio de ejercicio en la vida del contrato.

Para las ecuaciones obtenidas, Black y Scholes asumen que el precio de la opción es una función del precio de la acción y del tiempo a expirar. Sobre un intervalo de tiempo "corto", un portafolio con cobertura que incluya la acción común, la opción y un término corto de valores es construido para que el portafolio con porcentajes adecuados pueda eliminar todo "riesgo de mercado". Es decir, la componente de salto no se corralaciona con el mercado, entonces algún portafolio con un riesgo de mercado ("beta") igual a cero debe tener un rendimiento esperado igual a la tasa libre de riesgo. Aquí, una condición de equilibrio es establecida entre el rendimiento esperado en la opción, el rendimiento esperado de la acción y la tasa libre de riego.

Capital Asset Pricing Model

Definición 4 *Dado un portafolio con cobertura (o eficiente), entonces la gráfica de la ecuación 21 es llamada* ***capital market line****. Esta línea muestra la*

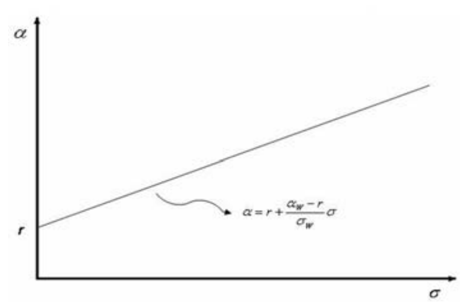

Figura 1: Capital market line.

relación entre el rendimiento esperado y el riesgo del rendimiento (medido por la desviación estándar) para el portafolio con cobertura. Donde la pendiente del capital market line es $(\alpha_W - r)/\sigma_W$ y este valor es frecuentemente llamado ***el precio de riesgo****.*

Supuesto 6 *Los inversionistas pueden escoger entre portafolios, en base a el rendimiento esperado y la varianza.*

Supuesto 7 *Todos los inversionistas estan de acuerdo con respecto a el horizonte y distribución de los rendimientos del valor.*

Y es válido el supuesto 2.

Teorema 3 *CAPM Si P^* es un portafolio con cobertura, entonces el rendimiento esperado α satisface*

$$\alpha - r = \beta(\alpha_W - r) \tag{25}$$

donde

$$\beta = \frac{\sigma}{\sigma_W} \tag{26}$$

Demostración

Para algún ω, considérese el portafolio P^* formado por una porción ω invertida en acciones y una porción $(1-\omega)$ invertida en opciones. (Notése que no se restringe a que $\omega \in [0,1]$, $\omega < 0$, que correspondería a pedir prestado a la tasa libre de riesgo r). El rendimiento esperado y la desviación estándar de este portafolio son

$$\alpha_{P^*} = \omega\alpha + (1-\omega)\alpha_W$$

$$\sigma_{P^*} = \left[\omega^2\sigma^2 + 2\omega(1-\omega)\sigma_W\sigma + (1-\omega)^2\sigma_W^2\right]^{1/2}$$

Calculando las parciales de α_{P^*} y σ_{P^*} y evaluandolas cuando $\omega = 0$

$$\frac{\partial\alpha_{P^*}}{\partial\omega} = \alpha - \alpha_W$$

$$\frac{\partial\sigma_{P^*}}{\partial\omega} = \frac{\omega\sigma^2 + (1-2\omega)\sigma_W\sigma + (\omega-1)\sigma_W^2}{\sigma_{P^*}}$$

En consecuencia

$$\left.\frac{\partial\sigma_{P^*}}{\partial\omega}\right|_{\omega=0} = \frac{\sigma\sigma_W - \sigma_W^2}{\sigma_W}$$

Por otro lado

$$\frac{\partial\alpha_{P^*}}{\partial\sigma_{P^*}} = \frac{(\partial\alpha_{P^*})/(\partial\omega)}{(\partial\sigma_{P^*})/(\partial\omega)}$$

Obteniéndose

$$\left.\frac{\partial\alpha_{P^*}}{\partial\sigma_{P^*}}\right|_{\omega=0} = \frac{(\alpha-\alpha_W)\sigma_W}{\sigma\sigma_W - \sigma_W^2}$$

Esta pendiente debe equivaler a la pendiente del capital market line. Es decir,

$$\frac{(\alpha-\alpha_W)\sigma_W}{\sigma\sigma_W - \sigma_W^2} = \frac{\alpha_W - r}{\sigma_W}$$

$$(\alpha-\alpha_W)\sigma_W = (\alpha_W - r)(\sigma - \sigma_W)$$

$$\alpha - \alpha_W = (\alpha_W - r)(\frac{\sigma}{\sigma_W} - 1)$$

$$\alpha = r + \frac{\sigma}{\sigma_W}(\alpha_W - r)$$

Por lo tanto

$$\alpha = r + \beta(\alpha_W - r)$$

□

La β de la acción es todo lo que se necesita conocer sobre las características del riesgo de la acción para usar el método del CAPM. El valor $\alpha - r$ es el término del **rendimiento esperado excesivo** de la acción; es la cantidad por la cual el rendimiento esperado excede a la tasa libre de riesgo. Asimismo, $\alpha_W - r$ es el exceso del rendimiento esperado de la opción.

0.1. El CAPM como modelo de valuación

El CAPM es un modelo de evaluación. Sin embargo la ecuación del CAPM estándar no contiene los precios explícitamente. Para ver que el CAPM es un modelo de valuación se debe partir de la definición de rendimiento.

Supóngase que una acción es adquirida al precio X y vendida al precio E. La tasa de rendimiento es entonces $\alpha = (E - X)/X$. Donde X es conocida y E es aleatoria. Sustituyéndose la tasa de rendimiento en el modelo CAPM, se tiene que

$$\frac{E - X}{X} = r + \beta(\alpha_w - r)$$

o

$$X = \frac{E}{1 + r + \beta(\alpha_w - r)}$$

Esta es la forma de valuación del CAPM, donde X es el precio de la acción; E es el precio de venta; y β es la beta de la acción.

Aplicación

A continuación se proporciona un ejemplo con la información expuestas hasta ahora.

Un banco puede definirse como una institución que posee recursos financieros y financia sus adquisiciones por equities y emitiendo depósitos que son completamente seguros. Desde el punto de vista de los depositantes, cada banco es un perfecto sustituto de cualquier otro banco, y por consiguiente $R + s$ debe ser el mismo para todos los bancos donde R es la tasa de interés en los depósitos y s es la tasa pagada en los servicios. Se asume que los depósitos son garantizados por el gobierno o una de sus agencias.

Para desarrollar el modelo se necesitan de los supuestos (2), (3), (4), (5) y de los siguientes supuestos:

Supuesto 8 *La dinámica para el valor de un recurso bancario dado, V, puede ser descrito por una ecuación diferencial estocástica*

$$\begin{aligned} dV &= [\alpha V - (R + s)D]\,dt + dD + \sigma dZ \\ &= [\alpha V - (R + s - g)D]\,dt + \sigma dZ \qquad V > 0 \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\quad\; V = 0 \end{aligned}$$

donde α el rendimiento esperado instantáneo en el recurso; σ^2 es la varianza instantánea y es constante; dZ es un proceso de Wiener.

Supuesto 9 *El gobierno o una de sus agencias cobra al banco una prima para asegurar todos los depósitos del banco a perpetuidad con tal de que el banco sea solvente (es decir, $V > D$). Para los propositos de este modelo, la solvencia del banco es determinada por auditorías [5]. Si, una auditoria encuentra al banco ser solvente , entonces el banco continua operando normalmente. Sin embargo, si el banco se ha encontrado insolvente (es decir, $V < D$), entonces el gobierno líquida al banco y paga los depositos.*

Supuesto 10 *Las auditorias son producidas aleatoriamente, donde el evento de una auditoria se distribuye Poisson con parámetro λ, y auditorias sucesivas son independientes e idénticamente distribuidas.*

Supuesto 11 *Existe un costo por auditoria, cual puede ser descrito por una función homogénea de primer grado $C(V, D)$. Por consiguiente, el costo de auditoria por pesos de los depositos, C/D, puede escribirse como $c(x)$ donde $x \equiv V/D$. Se asume que $c(x)$ es una función acotada y continua para toda $x \geq 0$.*

Cuando el gobierno asegura los depósitos de un banco, recibe una prima y asume una obligación de salidas de dinero en efectivo de futuros potenciales y actuales por los costos de vigilancia y por algún déficit entre los recursos y los depósitos en el caso de la liquidación del banco. Por conveniencia, el valor será deducido tratando a la salida de dinero como una entrada de dinero positiva, y por consiguiente el valor obtenido será positivo.

Se asume que el valor de la obligación puede escribirse como una función doblemente diferenciable $P(V, D)$ con las primeras derivadas continuas. El rendimiento de la obligación y sus cambios en el precio sobre un corto intervalo dependerá no sólo en los cambio de V y D, sino también en si o no hay una auditoria durante ese intervalo. Si una auditoria toma lugar, entonces hay un flujo de efectivo de $C(V, D)$. Además, si el banco es insolvente entonces hay un segundo flujo de dinero $D - V$ y la obligación del gobierno termina (es decir, $P = 0$). Si dR_P es la tasa de rendimiento de la obligación, entonces, usando el lema de Itô, se puede escribir dR_P como

$$\begin{aligned} PdR_P &= L\left[P(V,D)\right]dt + \sigma V \tfrac{\partial P}{\partial V} dZ \\ &\qquad\qquad\qquad \text{si no ocurre una auditoria} \\ &= L\left[P(V,D)\right]dt + \sigma V \tfrac{\partial P}{\partial V} dZ + C(V,D) \\ &\qquad\qquad\qquad \text{si ocurre una auditoria y } V > D \\ &= L\left[P(V,D)\right]dt + \sigma V \tfrac{\partial P}{\partial V} dZ + C(V,D) + D - V - P \\ &\qquad\qquad\qquad \text{si ocurre una auditoria y } V < D \end{aligned} \tag{27}$$

[5] Una excepción es cuando $V = 0$ y $g < R + s$. En este caso, el banco será incapaz de pagar el rendimiento prometido, $R + s$, en los depósitos. Esta “carencia” es asumida para causar una auditoria inmediata por el gobierno o sus agencias.

donde L es un operador definido por

$$L \equiv \frac{1}{2}\sigma^2 V^2 \frac{\partial^2}{\partial V^2} + \left[\alpha V - (R+s-g)D\right] \frac{\partial}{\partial V} + gD\frac{\partial}{\partial D}$$

Si el evento de una auditoria es independiente del rendimiento en el mercado de acciones, entonces de (27) el único origen de riesgo sistematico en el rendimiento en P es a través de su dependencia en V (es decir, el término dZ en (27)). Por consiguiente, la beta del rendimiento en P, β_P, puede ser escrita en términos de la beta del rendimiento en V, β_V, como

$$\beta_P = \frac{\sigma V \frac{\partial P}{\partial V}}{P} \frac{\beta_V}{\sigma} \tag{28}$$

Del supuesto (3) y (28), el equilibrio del rendimiento esperado en P, α_P, debe satisfacer

$$\alpha_P - r = \frac{V\frac{\partial P}{\partial V}(\alpha - r)}{P} \tag{29}$$

Puesto que el evento de una auditoria se distribuye Poisson, la probabilidad de una auditoria en el siguiente instante es λdt y la probablidad de que no ocurra la auditoria el $1-\lambda dt$. Tomando probabilidades en (27), se tiene

$$\begin{aligned} P\alpha_P &= L\left[P(V,D)\right] + \lambda C(V,D) && \text{para } V > D \\ &= L\left[P(V,D)\right] + \lambda\left[C(V,D) + D - V - P\right] && \text{para } V < D \end{aligned} \tag{30}$$

Combinando (29) y (30) y simplificando términos, se tiene que P debe satisfacer la ecuación deferencial parcial

$$\begin{aligned} &\tfrac{1}{2}\sigma^2 V^2 \tfrac{\partial^2 P}{\partial V^2} + \left[rV - (R+s-g)D\right] \tfrac{\partial P}{\partial V} + gD\tfrac{\partial P}{\partial D} - rP + \lambda C(V,D) = 0 \\ &\hspace{20em} \text{para } V > D \\ &\text{y} \\ &\tfrac{1}{2}\sigma^2 V^2 \tfrac{\partial^2 P}{\partial V^2} + \left[rV - (R+s-g)D\right] \tfrac{\partial P}{\partial V} + gD\tfrac{\partial P}{\partial D} - (r+\lambda)P \\ &\hspace{12em} + \lambda\left[C(V,D) + D - V\right] = 0 \\ &\hspace{20em} \text{para } V < D \end{aligned} \tag{31}$$

Las ecuaciones (31) pueden ser escritas en forma de ecuaciones diferenciales ordinarias por un cambio en la variable $x \equiv V/D$ y $p \equiv P/D$ donde x es la relación recurso-depósito y p es la obligación por cada peso de los depósitos. Si $p_1 \equiv p(x)$ para toda $x \geq 1$ y $p_2 \equiv p(x)$ para $0 \leq x \leq 1$, entonces de (31) se tiene

$$\begin{aligned} &\tfrac{1}{2}\sigma^2 x^2 p_1^{''} + \left[(r-g)x - (R+s-g)\right] p_1^{'} - (r-g)p_1 + \lambda C(x) = 0 \\ &\hspace{22em} \text{para } x \geq 1 \\ &\text{y} \\ &\tfrac{1}{2}\sigma^2 x^2 p_2^{''} + \left[(r-g)x - (R+s-g)\right] p_2^{'} - (r+\lambda-g)p_2 + \lambda\left[C(x) + 1 - x\right] = 0 \\ &\hspace{22em} \text{para } x \leq 1 \end{aligned} \tag{32}$$

Las condiciones de frontera para (32) son las siguientes

$$\begin{array}{lcll} p_1(1) & = & p_2(1) & \text{(continuidad de } p(x)) \\ p_1^{'}(1) & = & p_2^{'}(1) & \text{(continuidad de } p^{'}(x)) \\ p_2(0) & = & & \lambda\,[1+c(0)]\,/(r+\lambda-g) \\ p_1(x) & & & \text{esta acotada cuando } x\to\infty \end{array} \tag{33}$$

La solución homogénea general para (32) será asociada a funciones hipergeometricas. Sin embargo, en el caso particular donde la tasa de crecimiento en los depósitos agregados equivale al total del rendimiento en los depósitos (es decir, $g = R + s$), las soluciones son simples polinomios. En base a esta sustancial simplificación en las soluciones y adoptándose esto como un supuesto adicional, la solución completa a (32) puede ser escrita como

$$\begin{array}{lcll} p_1(x) & = & a_1 x + a_2 x^{-\delta} + Q_1(x) & \text{para } x \geq 1 \\ p_2(x) & = & b_1 x^k + b_2 x^{\xi} + \frac{\lambda}{\mu+\lambda} - x + Q_2(x) & \text{para } x \leq 1 \end{array} \tag{34}$$

donde varios de los parámetros y funciones son definidos como sigue

$$\begin{array}{lcl} k & \equiv & \frac{1}{2}\left\{1-\delta+\left[(1+\delta)^2+\gamma\right]^{1/2}\right\} > 1 \\ \xi & \equiv & 1-\delta-k < 0 \\ \mu & \equiv & r-g = r-(R+s) > 0 \\ \gamma & \equiv & 8\frac{\lambda}{\sigma^2} > 0 \\ \delta & \equiv & 2\frac{\mu}{\sigma^2} > 0 \\ Q_1(x) & \equiv & \frac{\lambda\delta}{\mu(1+\delta)}\left[x^{-\delta}\int^x y^{\delta-1}c(y)dy - x\int^x y^{-2}c(y)dy\right] \\ Q_2(x) & \equiv & \frac{\lambda\delta}{\mu(k-\xi)}\left[x^{\xi}\int^x y^{-\xi-1}c(y)dy - x^k\int^x y^{-k-1}c(y)dy\right] \end{array} \tag{35}$$

y $a_1,\ a_2,\ b_1$ y b_2 son constantes para ser escogidas tal que satisfagan las condiciones de frontera (33). Las soluciones en (34) dependen en la forma de la función de costos $c(x)$. Se asumirá además el caso especial donde el costo de la auditoria por peso de los depósitos es constante (es decir, $c(x) = K$). En este caso especial, la solución a (34), incluyendo las condiciones de frontera (33), puede ser escrita como

$$\begin{array}{lcll} p_1(x) & = & \frac{\lambda K}{\mu} - \frac{1}{\mu(\mu+\lambda)(\delta+k)}\left[\mu^2(k-1)+\lambda(\lambda K k-\mu)\right]x^{-\delta} & x \geq 1 \\ p_2(x) & = & \frac{\lambda(K+1)}{\mu+\lambda} - x + \frac{1}{\delta+k}\left[1+\frac{\delta}{\mu(\mu+\lambda)}(\mu^2+\lambda^2 K)\right]x^k & x \geq 1 \end{array} \tag{36}$$

Analizando (36) muestra que, la obligación del gobierno por cada peso de los depósitos no es una función monótonamente decreciente de la relación recurso-depósito. La razón es que hay dos procedencias de la obligación:

a) la garantía de los depósitos es una función monótona decreciente, y

b) el costo de la auditoria que es una función monótona creciente.

Riesgo de crédito

En este capítulo se presenta una introducción a los conceptos de riesgo, reisgo de crédito; y el impacto que estos tienen en el mundo financieron. Además, se concluye con una una proposición la cual es una solución aproximada para la valuación de una opción de cambio que esta sujeta al riesgo de incumplimiento.

Aspectos generales del riesgo

Se considera al riesgo como la posibilidad de una pérdida financiera debida a la componente aleatoria en la volatilidad del mercado. El riesgo puede dividirse en dos tipos:

1. *Riesgo específico o no-sistemático*: es la componente asociada a un activo por separado.
2. *Riesgo no específico o sistemático*: es el asociado con fatores que afectan a todo el mercado como un todo.

Por otro lado, crédito es el término utilizado para referirse a las transaciones que implican una transferencia de dinero que debe devolverse transcurrido cierto tiempo. Por tanto, el que transfiere el dinero se convierte en acreedor y el que lo recibe en deudor; los términos crédito y deuda reflejan pues una misma transacción desde dos puntos de vista opuestos.

De lo anterior se puede concluir que el riesgo de crédito es el asociado a la posibilidad de que la contraparte sea incapaz de cumplir los créditos pactados en el contrato.[1] Esto debido a situaciones financieras extremas que llevan a no tener solvencia causando así una gran pérdida económica.

El riesgo de crédito por incumplimiento, ya sea en un contrato en general o en un crédito, se considera como un factor crucial para las valuación de instrumentos de deuda.

Las siguientes son algunas razones para considerar el riesgo de crédito de una forma más estricta:

i) El manejo del riesgo de crédito en los bancos y otras instituciones empieza a ser inapropiado. Por lo que se desarrollan productos derivados para cubrirse de la inestabilidad económica como los swaps y las opciones, que contienen una gran parte del riesgo de crédito.

[1] Hay diferencia entre el riesgo de incumplimiento y el riesgo de crédito. El riesgo de incumplimiento es el riesgo que considera la posibilidad de que la contraparte falle con el compromiso. El riesgo de crédito tiene que ver con el incumplimiento del bien subyacente.

ii) El mercado fuera de mostrador constituye la mayor parte del total de productos derivados. La emisión de instrumento fuera de mostrador se realiza usualmente sin garantía y en la mayoría de los casos es insegura la reclamación.

iii) Históricamente las tasas de incumplimiento indican un elevado incremento. Las posibilidades tan altas en este riesgo, las inversiones en productos derivados exigen tomar en cuenta el riesgo de crédito.

Derivados de crédito con riesgo de incumplimiento

En esta sección se hace un análisis del marco desarrollado por Heath, Jarrow y Morton en 1992 sobre la tasa de interés.

Sea un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbf{P})$. El intervalo de intercambio es continuo en $t \in [0, T]$. Se asume que el mercado consiste de n valores tal que el vector $S(t) = \left(S_t^1, \ldots, S_t^n\right)$ es un proceso estocástico $\mathcal{F}_t$-adaptado en $\mathbb{R}_+^n$ que modela los precios de los valores. Además se asume que se tiene un mercado perfecto donde el intercambio toma lugar continuamente y no hay fricciones.

Sea $W(t)$ un movimiento Browniano en $\mathbb{R}^d$ definido en el espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbf{P})$. Además se asume que hay una medida de probabilidad $\mathbf{Q}$ equivalente a $\mathbf{P}$ tal que, bajo la medida $\mathbf{Q}$ los precios de las acciones se comportan como una cuenta de mercado de dinero $B(t)$ y son $\mathbf{Q}$-martingalas.

Definición 1 *Considérese el intervalo de tiempo $\mathcal{T} = [0, T']$. El precio al tiempo t de un bono* [2] *con fecha de vencimiento al tiempo T es escrito como $P(t, T)$ para alguna $t,\ T\ \in \mathcal{T}$ tal que $t \leq T$.*

La tasa forward instantánea $f(t, T)$ esta dada por

$$f(t,T) = -\tfrac{1}{P(t,T)} \tfrac{\partial}{\partial T} P(t,T) = -\tfrac{\partial}{\partial T} \ln P(t,T) \quad \forall\ t \leq T$$

donde $f(t, T)$ es la tasa de rendimiento disponible al tiempo t para una inversión instantánea al tiempo T.

Alternativamente, $P(t, T)$ puede ser expresada en términos de la tasa forward

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right)$$

donde $f(t, T)$ está en $(\mathcal{L}^1)$ [3] para asegurar la integrabilidad.

[2]Donde por bono se entenderá un bono de cupón cero.

[3]Donde $\mathcal{L}^{(*)}$ define una clase de procesos. Consultar A.2 en el apéndice.

La tasa a corto plazo es un caso especial de la tasa forward, es decir

$$r(t) = f(t,t)$$

La cuenta de mercado de dinero, la cuál sirve como numerario, se define en la economía con la tasa de interés determinista, es decir, $dB(t) = r(t)B(t)dt$ con $B(0) = 1$. Esta dada por

$$B(t) = \exp\left(\int_t^T r(s)ds\right) \tag{1}$$

La estructura de la tasa forward es asumida a seguir un proceso de Itô tal que para alguna $T \in [0, T']$ fija y $\alpha(t) \in \mathcal{L}^1$ y $\sigma(t) \in \mathcal{L}^2$,

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \sigma(s,T)dW(s)$$

$\forall\ T$ tal que $0 \leq t \leq T$, donde $W(t)$ toma valores en $\mathbb{R}^d$ y esta definido en un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbf{P})$. Los procesos $\alpha(t,T)$ y $\sigma(t,T)$ toman valores en $\mathbb{R}$ y $\mathbb{R}^d$, respectivamente. Además se requiere que $\alpha(t,T) \in \mathcal{L}^{1,1}$ y $\sigma(t,T) \in (\mathcal{L}^{2,1})^d$.

Por lo anterior se observa que la tasa a corto plazo obedece

$$r(t) = f(0,t) + \int_0^t \alpha(s,t)ds + \int_0^t \sigma(s,t)dW(s) \qquad \forall\ t \in [0, T'] \tag{2}$$

Para que el proceso de la cuenta de mercado de dinero este bien definida, se requiere que $f(0,t) \in \mathcal{L}^1$, además de las condiciones de regularidad para α y σ arriba especificadas. Entonces de (1) y (2) se tiene que

$$B(t) = \exp\left(\int_0^t f(0,u)du + \int_0^t \int_s^t \alpha(s,u)duds + \int_0^t \int_s^t \sigma(s,u)dudW(s)\right) \tag{3}$$

Similarmente, el precio del bono puede ser escrito en términos de la tasa forward.

$$P(t,T) = \exp\left(-\int_t^T f(0,s)ds - \int_0^t \int_t^T \alpha(s,u)duds - \int_0^t \int_t^T \sigma(s,u)dudW(s)\right) \tag{4}$$

Proposición 1 *El proceso del precio del bono en términos del dinero de mercado, $Z(t,T) = B(t)^{-1}P(t,T)$ esta dado por*

$$\ln Z(t,T) = -\int_0^T f(0,s)ds - \int_0^t \int_s^T \alpha(s,u)duds - \int_0^t \int_s^T \sigma(s,u)dudW(s)$$

o en su respectiva forma de ecuación diferencial estocástica

$$\frac{dZ(t,T)}{Z(t,T)} = -\left(\int_t^T \alpha(t,u)dudt - \frac{1}{2}\int_t^T ||\sigma(t,u)||^2 dudt + \int_t^T \sigma(t,u)dudW(t)\right)$$

Demostración

El logaritmo del precio del bono en términos del dinero en efectivo puede ser escrito como $\ln Z(t,T) = \ln P(t,T) - \ln B(t)$. Por las expresiones (3) y (4) se tiene que

$$\begin{aligned}\ln Z(t,T) = &-\int_t^T f(0,s)ds - \int_0^t \int_t^T \alpha(s,u)duds - \int_0^t \int_t^T \sigma(s,u)dudW(s) \\ &-\int_0^t f(0,u)du - \int_0^t \int_s^t \alpha(s,u)duds - \int_0^t \int_s^t \sigma(s,u)dudW(s)\end{aligned}$$

Claramente, se puede ver que combinando las integrales se obtiene la proposición. □

Definición 2 *Para simplificar la notación, se usará*

$$\acute{\alpha}(t,T) = \int_t^T \alpha(t,s)ds \quad y \quad \acute{\sigma}(t,T) = \int_t^T \sigma(t,s)ds$$

Proposición 2 *Si el sistema de ecuación lineal*

$$-\acute{\alpha}(t,T) + \frac{1}{2}||\acute{\sigma}(t,T)||^2 = \acute{\sigma}(t,T) \cdot \gamma(t) \tag{5}$$

admite una única solución para el proceso $\gamma(t) \in \mathcal{L}^2$ que toma valores en $\mathbb{R}^d$ tal que la condición de Novikov [4] *es satisfecha, entonces existe una medida de martingala equivalente para Z.*

Demostración

Considérese a la ecuación diferencial estocástica deducida en la Proposición 1 para $Z(t,T)$. Por la simplificación de la notación introducida se puede escribir a la ecuación diferencial estocástica para $Z(t,T)$ como

$$dZ(t,T) = -Z(t,T)\left(\acute{\alpha}(t,T)dt + \acute{\sigma}(t,T)dW(t) - \frac{1}{2}||\acute{\sigma}(t,T)||^2 dt\right)$$

Sí $\gamma(t)$ es una solución única a la expresión (5), entonces

$$dZ(t,T) = -Z(t,T)\left(-\acute{\sigma}(t,T)\gamma(t)dt + \acute{\sigma}(t,T)dW(t)\right)$$

Se define una medida de probabilidad equivalente[5] tal que

$$\frac{d\mathbf{Q}}{d\mathbf{P}} = \exp\left(\int_0^T \gamma(s)dW(s) - \frac{1}{2}\int_0^T ||\gamma(s)||^2 ds\right)$$

[4] Consultar la sección A.7 del apéndice.
[5] Para saber más al respecto consultar la sección A.4 del apéndice.

Por el Teorema de Girsanov [6], $\tilde{W}(t) = W(t) - \int_0^t \gamma(s)ds$ es un movimiento Browniano estándar bajo $\mathbf{Q}$. En consecuencia,

$$dZ(t,T) = -Z(t,T)\left(-\acute{\sigma}(t,T)\gamma(t)dt + \acute{\sigma}(t,T)\cdot(d\tilde{W}(t) + \gamma(t)dt)\right).$$

Claramente, esto es $dZ(t,T) = -Z(t,T)\acute{\sigma}(t,T)\cdot d\tilde{W}(t)$ la cual es una martingala bajo $\mathbf{Q}$ si la condición de Novikov es satisfecha. □

Proposición 3 *Sí para $T \in [0, T']$ arbitraría,*

$$\alpha(t,T) = \sigma(t,T)\cdot\acute{\sigma}(t,T), \qquad \forall\ t \in [0,T'] \quad \textit{tal que} \quad t \leq T$$

bajo la medida de probabilidad equivalente $\mathbf{Q}$ y tal que $\alpha(t,T) \in \mathcal{L}^{1,1}$, entonces las dinámicas de los precios de los bonos son libres de arbitrage.

Demostración

Por la Proposición 2, existe una única medida equivalente de martingala si $-\acute{\alpha}(t,T) + \frac{1}{2}||\acute{\sigma}(t,T)||^2 = \acute{\sigma}(t,T)\cdot\gamma(t)$ tiene una única solución $\gamma(t)$ en $\mathbb{R}^d$. Por consiguiente, de la demostración de la Proposición 2, la tendencia del proceso $Z(t)$ es $-\acute{\alpha}(t,T)dt + \frac{1}{2}||\acute{\sigma}(t,T)||^2dt = \acute{\sigma}(t,T)\cdot\gamma(t)dt$. Por Girsanov, cambiando la medida de probabilidad de $\mathbf{P}$ a $\mathbf{Q}$ decrece la tendencia para $\gamma(t)dt$. Entonces, el término $\gamma(t)$ se cancela tal que el coeficiente de tendencia $-\acute{\alpha}(t,T)dt + \frac{1}{2}||\acute{\sigma}(t,T)||^2dt = 0$. Por lo tanto, la diferenciación da $\alpha(t,T) = \sigma(t,T)\cdot\acute{\sigma}(t,T)$. □

En base a lo anterior, se pueden determinar los procesos y las ecuaciones diferenciales estocásticas de la tasa de interés y los procesos del bono. El proceso de la tasa forward bajo $\mathbf{Q}$ se sigue inmediatamente de la Proposición 3.

$$f(t,T) = f(0,T) + \int_0^t \sigma(s,T)\cdot\acute{\sigma}(s,T)ds + \int_0^t \sigma(s,T)\cdot d\tilde{W}(s) \tag{6}$$

$\forall\ t$ tal que $0 \leq t \leq T$. En su forma diferencial,

$$df(t,T) = \sigma(t,T)\cdot\acute{\sigma}(t,T)dt + \sigma(t,T)\cdot d\tilde{W}(t)$$

Para determinar el proceso del precio del bono bajo $\mathbf{Q}$, se necesitan hacer algunas transformaciones. La definición del bono en términos de la tasa forward esta dada por la expresión (4). Ahora considérese la siguiente descomposición de la integral del logaritmo del precio del bono.

$$\begin{aligned}\ln P(t,T) = &-\int_0^T f(0,u)du - \int_0^t\int_s^T \alpha(s,u)duds - \int_0^t\int_s^T \sigma(s,u)dudW(s)\\ &+\int_0^t f(0,u)du + \int_0^t\int_s^t \alpha(s,u)duds + \int_0^t\int_s^t \sigma(s,u)dudW(s),\end{aligned}$$

[6] El enunciado de este teorema se encuentra en la sección A.4 del apéndice.

o en su forma simplificada

$$\ln P(t,T) = -\int_0^T f(0,u)du - \int_0^t \acute{\alpha}(s,T)ds - \int_0^t \acute{\sigma}(s,T)dW(s)$$
$$+\int_0^t f(0,u)du + \int_0^t \acute{\alpha}(s,t)ds + \int_0^t \acute{\sigma}(s,t)dW(s),$$

Por la expresión (1) y (3), el sumando de la expresión anterior es igual a $\int_0^t r(s)ds$. Además, se tiene que $-\int_0^T f(0,u)du = \ln P(0,T)$.

Ahora se introduce la condición de arbitrage de la Proposición 2 y se hace un cambio de medida de probabilidad de $\mathbf{P}$ a $\mathbf{Q}$ usando la sustitución $dW(t) = d\tilde{W}(t) + \gamma(t)dt$. Con lo cual se obtiene

$$\ln P(t,T) = \ln P(0,T) - \frac{1}{2}\int_0^t ||\acute{\sigma}(s,T)||^2 ds - \int_0^t \acute{\sigma}(s,T)d\tilde{W}(s)$$
$$+\int_0^t f(0,u)du + \frac{1}{2}\int_0^t ||\acute{\sigma}(s,t)||^2 ds + \int_0^t \acute{\sigma}(s,t)d\tilde{W}(s),$$

Puesto que $r(t) = f(t,t)$ y $f(t,T)$ bajo $\mathbf{Q}$ esta dado por (6), $r(t)$ bajo $\mathbf{Q}$ puede ser escrito como

$$r(t) = f(0,t) + \int_0^t \sigma(s,t)\cdot \acute{\sigma}(s,t)ds + \int_0^t \sigma(s,t)\cdot d\tilde{W}(s) \tag{7}$$

Integrando la expresión anterior se obtiene

$$\int_0^t r(s)ds = \int_0^t f(0,u)du + \frac{1}{2}\int_0^t ||\acute{\sigma}(s,t)||^2 ds + \int_0^t \acute{\sigma}(s,t)\cdot d\tilde{W}(s)$$

Por consiguiente, el proceso del precio del bono bajo la medida $\mathbf{Q}$ esta dado por

$$P(t,T) = P(0,T)\exp\left(\int_0^t r(s)ds - \frac{1}{2}\int_0^t ||\acute{\sigma}(s,T)||^2 ds - \int_0^t \acute{\sigma}(s,T)d\tilde{W}(s)\right) \tag{8}$$

Aplicando la formula de Itô se obtiene la correspondiente ecuación diferencial estocástica

$$dP(t,T) = P(t,T)\left(r(t)dt - \acute{\sigma}(t,T)d\tilde{W}(t)\right)$$

Como se esperaba, la tendencia del precio del bono es la tasa corta bajo la medida de probabilidad $\mathbf{Q}$.

Medida forward

Hasta ahora, el dinero en efectivo ha sido usualmente el mercado de dinero $B(t)$. La medida de probabilidad $\mathbf{Q}$ equivalente asociada es llamada la medida de riesgo neutral. Bajo $\mathbf{Q}$, el precio de las acciones son martingalas en términos del dinero en efectivo $B(t)$, es decir,

$$I\!E_{\mathbf{Q}}\left[B^{-1}(T)S(T)|\mathcal{F}_t\right] = B^{-1}(t)S(t),$$

para cualquier t arbitraria tal que $t \leq T$.

Aquí el precio del bono $P(t,T)$ es usado como dinero en efectivo. Se muestra que existe una medida de probabilidad equivalente tal que los precios de las acciones son martingalas en términos del nuevo dinero en efectivo. Si el nuevo dinero en efectivo es un bono cupón cero, esta medida de probabilidad $\mathbf{F}$ es algunas veces llamado *forward neutral* ya que los precios de las acciones medidos en términos del precio del bono tienen la interpretación financiera de los precios forward.

Si existe una medida de probabilidad equivalente tal que los precios medidos en términos de los precios de los bonos son martingalas, entonces

$$I\!E_{\mathbf{F}}\left[P(T,T)^{-1}S(T)|\mathcal{F}_t\right] = I\!E_{\mathbf{F}}\left[S(T)|\mathcal{F}_t\right] = S(t)$$

para cualquier t arbitraria tal que $t \leq T$. Es decir, los precios forward son estimados imparciales de los precios futuros bajo $\mathbf{F}$.

El precio de un título que paga $X(T)$ dependiendo de la acción $S(T)$ al tiempo T puede ser escrito

$$X(t) = P(t,T)I\!E_{\mathbf{F}}\left[P(T,T)^{-1}X_T(S(T,T))|\mathcal{F}_t\right]$$

ya que $P(T,T) = 1$ se tiene que

$$X(t) = P(t,T)I\!E_{\mathbf{F}}\left[X_T(S(T,T))|\mathcal{F}_t\right].$$

Definición 3 *El precio de un bono $P(t,\tau)$ en términos del bono (dinero en efectivo) está dado por*

$$F(t,T,\tau) = \frac{P(t,\tau)}{P(t,T)}$$

donde $F(t,T,\tau)$ es llamado el precio forward de $P(t,\tau)$.

Definición 4 *Se define una medida de probabilidad $\boldsymbol{F}$ equivalente a $\boldsymbol{Q}$ tal que*

$$\frac{d\boldsymbol{F}}{d\boldsymbol{Q}} = \frac{P(T,T)P(0,T)^{-1}}{B(T)B(0)^{-1}} = \frac{1}{P(0,T)B(T)}$$

donde $\boldsymbol{F}$ es conocida como la medida forward. Ahora se define a $\zeta(t)$ como la derivada de Radon-Nikodým tal que

$$\zeta(t) = \frac{P(t,T)}{P(0,T)B(t)}.$$

Considérese las expresiones (8), (1) y (7). Dados esos procesos, la derivada de Radon-Nikodým para cambiar medida de **Q** a **F** puede ser determinada por una sustitución de $P(0,T)$, $P(t,T)$ y $B(t)$.

$$\zeta(t) = \frac{P(t,T)}{P(0,T)B(t)} = \exp\left(-\frac{1}{2}\int_0^t ||\sigma(s,T)||^2 ds - \int_0^t \acute{\sigma}(s,T)d\tilde{W}(s)\right) \quad (9)$$

por la formula de Itô, se obtiene la correspondiente ecuación diferencial estocástica,

$$d\zeta(t) = -\zeta(t)\acute{\sigma}(t,T)\cdot d\tilde{W}(t).$$

Proposición 4 *El precio de un bono en términos del bono (dinero en efectivo) $F(t,T,\tau)$ es una martingala bajo **F** si satisface la condición de Novikov.*

Demostración

Consultar Manuel Ammann (2001). □

Proposición 5 *Bajo la medida **F**, la tasa forward satisface la ecuación diferencial estocástica*

$$df(t,T) = \sigma(t,T)\cdot d\tilde{W}(t)$$

Demostración

De la expresión (6), la ecuación diferencial estocástica para la tasa forward bajo **Q** es

$$df(t,T) = \sigma(t,T)\cdot d\tilde{W}(t) + \sigma(t,T)\cdot \int_t^T \sigma(t,u)dudt$$

Por el teorema de Girnasov con $d\tilde{W}(t) = d\hat{W}(t) - \int_t^T \sigma(t,u)dudt$, se obtiene la proposición. □

Por otro lado, de la Definición 3, el precio forward es el precio del bono en términos de un bono numéraire, es decir,

$$F(t,T,\tau) = \frac{P(t,\tau)}{P(t,T)}$$

además, la Proposición 4 establece que, bajo una medida **F** martingala definida en (9), el precio forward del bono es una martingala. Entonces la ecuación diferencial estocástica está dada por

$$\frac{dF(t,T,\tau)}{F(t,T,\tau)} = -\int_T^\tau \sigma(t,u)du\cdot d\hat{W}(t)$$

para simplificar la notación, se define a

$$\beta(t,T,\tau) = -\int_T^\tau \sigma(t,u)du. \quad (10)$$

Se requiere que β sea una función acotada determinista.

El proceso del precio de una acción bajo la medida $\mathbf{Q}$ es

$$S(t) = S(0)\exp\left(\int_0^t r(s)ds - \frac{1}{2}\int_0^t ||\sigma_S(s)||^2 ds + \int_0^t \sigma_S(s)d\tilde{W}(s)\right).$$

El precio forward de una acción está dado por

$$S_F(t,T) = \frac{S(t)}{P(t,T)}.$$

La ecuación diferencial estocástica de S_F bajo $\mathbf{F}$ esta dada de la siguiente forma. Bajo $\mathbf{Q}$ las dinámicas de S_F son

$$\begin{aligned}\frac{dS_F(t,T)}{S_F(t,T)} &= \left(\sigma_S(t) + \int_t^T \sigma(t,u)du\right)\cdot d\tilde{W}(t) \\ &\quad + \int_t^T ||\sigma(t,u)||^2 dudt + \sigma_S(t)\cdot \int_t^T \sigma(t,u)dudt\end{aligned}$$

Cambiándose la medida de $\mathbf{Q}$ a $\mathbf{F}$ implica, que por del teorema de Girsanov, se tiene

$$\begin{aligned}\frac{dS_F(t,T)}{S_F(t,T)} &= \left(\sigma_S(t) + \int_t^T \sigma(t,u)du\right)\cdot\left(d\hat{W}(t) - \int_t^T \sigma(t,u)dudt\right) \\ &\quad + \int_t^T ||\sigma(t,u)||^2 dudt + \sigma_S(t)\cdot \int_t^T \sigma(t,u)dudt \\ &= \left(\sigma_S(t) + \int_t^T \sigma(t,u)du\right)\cdot d\hat{W}(t)\end{aligned}$$

Como en (10), se denotará al vector volatilidad del precio de la acción relativo al dinero en efectivo $P(t,T)$ por $\beta_S(t,T)$, como

$$\beta_S(t,T) = \sigma_S(t) + \int_t^T \sigma(t,u)du \tag{11}$$

tal que

$$\frac{dS_F(t,T)}{S_F(t,T)} = \beta_S(t,T)\cdot d\hat{W}(t).$$

El proceso para $V(t)$ es definido análogamente, con parámetro de volatilidad β_V bajo la medida forward $\mathbf{F}$.

Las expresiones (10) y (11) son estrictamente análogas, lo cual se puede verificar por la separación de la integral en (10), es decir,

$$\beta(t,T,\tau) = -\int_t^\tau \sigma(t,u)du + \int_t^T \sigma(t,u)du.$$

De hecho, β denota la diferencia entre el precio de la acción y la volatilidad del dinero en efectivo. Por consiguiente se tiene que las funciones de volatilidad forward, para los bonos y las acciones, son respectivamente

$$\begin{aligned} \beta(t,T,\tau) &= b(t,\tau) - b(t,T) \\ \beta(t,T) &= \sigma_S(t) - b(t,T) \end{aligned} \tag{12}$$

con $b(t,T) = -\acute{\sigma}(t,T) = \int_t^T \sigma(t,s)ds$. Las expresiones anteriores se aplican análogamente para el proceso del valor de la empresa. El valor de la empresa forward se escribe $V_F(t,T)$.

El movimiento Browniano $\hat{W}(t)$ es un proceso en $\mathbb{R}^d$. Puesto que $\hat{W}(t)$ debe ser interpretado como un vector de independientes movimientos Brownianos de dimensión d, la varianza forward de la acción y el valor de la empresa están dados por

$$\nu_i^2(t,T) = \int_t^T ||\beta_i(s,T)||^2 ds, \qquad \forall\ i \in S, V.$$

Similarmente, la covarianza y correlación son definidas como

$$\phi(t,T) = \int_t^T \beta_S(s,T)\cdot \beta_V(s,T)ds$$

$$\rho(t,T) = \int_t^T \frac{\beta_S(s,T)\cdot \beta_V(s,T)}{||\beta_S(s,T)||||\beta_V(s,T)||}ds.$$

Precio de una opción de cambio con riesgo de incumplimiento

La siguiente proposición da un acercamiento a la solución para el precio de una opción de cambio que esta sujeta a riesgo de incumplimiento por parte de la contraparte. Esté modelo de riesgo de crédito está deacuerdo con el modelo de Merton (1974). Sea $X(T)$ la cantidad prometida de pago por la contraparte de la empresa al tiempo T. Si la empresa permanece solvente a lo largo de la vida del crédito, se pagará $X(T)$ al tiempo T. Se asume que X es un título tipo "Europeo", es decir, no puede ser reclamado antes del tiempo T.

Si la empresa sufre una quiebra algun dia denotado por t antes de la fecha de vencimiento T, el propietario del título no podra recibir $X(T)$, sino sólo una porción de ésta. La cantidad del título del propietario recibida es dependiente del valor de los recursos de la empresa relativos a sus obligaciones. Se denotará el valor de los recursos por $V(t)$ y el valor de las obligaciones por $D(t)$. Puesto que $\delta(T) = V(T)/D(T)$ denota el valor recuperado, es comúnmente llamado "*tasa de recuperación*". Sin embargo, las obligaciones de la contraparte se asumen ser constantes.

Los procesos de los precios del bono y la acción son especificados en la sección anterior. El valor de las acciones derivadas de la contraparte se asumen tener el siguiente comportamiento,

$$V(t) = V(0)\exp\left(\int_0^t r(s)ds - \frac{1}{2}\int_0^t ||\sigma_V(s)||^2 ds + \int_0^t \sigma_V(s)d\tilde{W}(s)\right)$$

bajo la medida de probabilidad $\mathbf{Q}$.

Proposición 6 *El precio $X(t)$ de una opción call con ganacia neta prometida $X(T) = \text{máx}\{S_1(T) - S_2(T), 0\}$ y ganancia neta real*

$$X(T) = \theta(T)\,\text{máx}\{S_1(T) - S_2(T), 0\}$$

donde $\theta(T) = V(T/D$ en caso de quiebra esta dado por

$$\begin{aligned} X(T) &= S_1(t)N_2(a_1, a_2, \rho_1(t,T)) - S_2(t)N_2(b_1, b_2, \rho_2(t,T)) \\ &\quad + \frac{S_1(t)V(t)}{DP(t,T)}e^{\phi_{S_1V(t,T)}}N_2(c_1, c_2, -\rho_1(t,T)) \\ &\quad - \frac{V(t)K}{D}N_2(d_1, d_2, -\rho_2(t,T)) \end{aligned}$$

con los siguientes parámetros

$$\begin{aligned} a_1 &= \frac{In\frac{S_1(t)}{S_2(t)} + \frac{1}{2}\chi^2(t,T)}{\chi(t,T)} \\ a_2 &= \frac{InV(t) - InD - InP(t,T) + \phi_{S_1V}(t,T) - \frac{1}{2}\nu_V^2(t,T)}{\nu_V(t,T)} \\ b_1 &= \frac{In\frac{S_1(t)}{S_2(t)} - \frac{1}{2}\chi^2(t,T)}{\chi(t,T)} \\ b_2 &= \frac{InV(t) - InD - InP(t,T) + \phi_{S_2V}(t,T) - \frac{1}{2}\nu_V^2(t,T)}{\nu_V(t,T)} \\ c_1 &= \frac{In\frac{S_1(t)}{S_2(t)} + \frac{1}{2}\chi^2(t,T) + \phi_{S_1V}(t,T)}{\chi(t,T)} \\ c_2 &= -\frac{InV(t) - InD - InP(t,T) + \phi_{S_2V}(t,T) + \frac{1}{2}\nu_V^2(t,T)}{\nu_V(t,T)} \\ d_1 &= \frac{In\frac{S_1(t)}{S_2(t)} - \frac{1}{2}\chi^2(t,T) + \phi_{S_1V}(t,T)}{\chi(t,T)} \\ d_2 &= -\frac{InV(t) - InD - InP(t,T) + \phi_{S_2V}(t,T) + \frac{1}{2}\nu_V^2(t,T)}{\nu_V(t,T)} \end{aligned}$$

y

$$\begin{aligned}
\nu_i^2(t,T) &= \int_t^T ||\beta_i(s,T)||^2 ds \qquad \forall\ i \in \{S_1, S_2, V\} \\
\phi_{i,j}(t,T) &= \int_t^T \beta_i(s,T)\cdot\beta_j(s,T) ds \qquad \forall\ i,j \in \{S_1, S_2, V\} \\
\rho_i(t,T) &= \int_t^T \frac{\beta_{S_i}(s,T)\cdot\beta_V(s,T)}{||\beta_{S_i}(s,T)||||\beta_V(s,T)||} ds \qquad \forall\ i \in \{1,2\} \\
\chi^2(t,T) &= \int_t^T ||\beta_{S_1}(s,T) - \beta_{S_2}(s,T)||^2 ds
\end{aligned}$$

donde β_i*,* $\forall i \in \{S_1, S_2, V\}$*, es una función acotada y determinista.*

En un contexto de los derivados de crédito, $S_1(t)$[7] y $S_2(t)$ son interpretados como las variables que determinan el pago de los derivados de crédito mientras que $V(t)$ denota el valor de la acción de la contraparte de la empresa del contrato de derivado. Obviamente, sobre la interpretación y el uso de la forma de valuación no está restringida a los derivados de crédito. Puede ser usada para valuar alguna opción de cambio que está sujeta a una contraparte riesgosa dado que el proceso estocástico usado para $S_1(t)$ y $S_2(t)$ representa un modelo razonable para el proceso real del bien subyacente.

Demostración

El precio del título de cambio puede ser escrito como

$$X(t) = P(t,T)I\!E_{\mathbf{F}}\left[(S_1(T,T) - S_2(T,T))^+ (\mathbf{1}_{\{V_F(T,T)\geq D\}} + \theta_F(T,T)\mathbf{1}_{\{V_F(T,T)<D\}})|\mathcal{F}_t\right]$$

donde $\theta_F(T,T) = \frac{V_F(T,T)}{D}$. $\mathbf{F}$ denota la medida neutral forward.

Esta expresión puede ser dividida en cuatro términos tal que

$$X(t) = E_1 - E_2 + E_3 - E_4$$

donde

$$\begin{aligned}
E_1 &= P(t,T)I\!E_{\mathbf{F}}\left[S_1^F(T,T)\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right] \\
E_2 &= P(t,T)I\!E_{\mathbf{F}}\left[S_2^F(T,T)\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right] \\
E_3 &= P(t,T)I\!E_{\mathbf{F}}\left[S_1^F(T,T)\theta_F(T,T)\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)<D\}}|\mathcal{F}_t\right] \\
E_4 &= P(t,T)I\!E_{\mathbf{F}}\left[S_2^F(T,T)\theta_F(T,T)\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)<D\}}|\mathcal{F}_t\right]
\end{aligned}$$

[7] El precio de la acción es descrito por $dS_i(t) = S_i(t)(\alpha_i(t)dt - \sigma_i(t)dW(t))$, donde $S_i(t)$ y $\alpha_i(t)$ estan valuados en $\mathbb{R}$, $\sigma_i(t)$ en $\mathbb{R}^d$ para toda $i \in \{1, 2, \ldots, n\}$.

Bajo la medida neutral forward, la dinámica del precio forward de la acción S_i^F y V_F están dados por

$$\frac{dS_i^F(t,T)}{S_i^F(t,T)} = \beta_{S_i}(t,T)\cdot d\hat{W}(t), \qquad \frac{dV_F(t,T)}{S_F(t,T)} = \beta_V(t,T)\cdot d\hat{W}(t)$$

donde $\beta_{S_i}(t,T)$ y $\beta_V(t,T)$ están definidas en (12) como

$$\beta_{S_i}(t,T) = \sigma_{S_i}(t) + \int_t^T \sigma(t,u)du$$

adicionalmente, se define

$$\eta_i(t,T) = \int_t^T \beta_i(s,T)\cdot d\hat{W}(s), \qquad \forall\ i \in \{S_1, S_2, V\}$$

$S_F(T,T)$ y $V_F(T,T)$ pueden por consiguiente ser escritas

$$i_F(T,T) = i_F(t,T)\exp\left(\eta_i(t,T) - \frac{1}{2}\nu_i^2(t,T)\right).$$

También se tiene que $i_F(T,T) = i_T$ y $i_F(t,T) = i_t P(t,T)^{-1}$, por la definición de precio forward.

A continuación se evaluará la expresión E_1

$$E_1 = P(t,T)\mathbb{E}_{\mathbf{F}}\left[S_1^F(T,T)\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right].$$

Pero sustituyendo a $S_1^F(T,T)$, se obtiene

$$\begin{aligned} E_1 &= \mathbb{E}_{\mathbf{F}}\left[\frac{S_1(t)}{P(t,T)}\exp\left(\int_t^T \beta_{S_1}(s,T)d\hat{W}(s) - \frac{1}{2}\int_t^T ||\beta_{S_1}(s,T)||^2 ds\right)\right. \\ &\quad \left.\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right] P(t,T) \\ &= S_1(t)\mathbb{E}_{\mathbf{F}}\left[\exp\left(\eta_{S_1}(t,T) - \frac{1}{2}\nu_{S_1}^2(t,T)\right)\right. \\ &\quad \left.\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right] \end{aligned} \tag{13}$$

Ahora se introduce una nueva medida $\dot{\mathbf{F}}$ definida por

$$\zeta(T) = \frac{d\dot{\mathbf{F}}}{d\mathbf{F}} = \exp\left(\int_0^T \beta_{S_1}(s,T)d\hat{W}(s) - \frac{1}{2}\int_0^T ||\beta_{S_1}(s,T)||^2 ds\right), \tag{14}$$

con $\zeta(t) = \mathbb{E}_{\mathbf{F}}\left[\zeta(T)|\mathcal{F}_t\right]$, $\forall\ t \in [0,T]$. La expresión (13) es igual a

$$E_1 = \zeta(t)^{-1}\mathbb{E}_{\mathbf{F}}\left[S_1(t)\zeta(T)\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right].$$

Por la regla de Bayes, esto es igual a

$$E_1 = S_1(t) I\!E_{\dot{\mathbf{F}}} \left[\mathbf{1}_{\left\{S_1^F(T,T) > S_2^F(T,T)\right\}} \mathbf{1}_{\{V_F(T,T) \geq D\}} | \mathcal{F}_t \right].$$

Donde $S_i^F(T,T)$ y $V_F(T,T)$ son funciones de $\eta_{S_i}(t,T)$ y $\eta_V(t,T)$, respectivamente. Por la definición de movimiento Browniano, $\eta_i(t,T),\ \forall\ i \in \{S_1, S_2, V\}$, tienen ley $\mathbf{F}$ $N(0, \nu_i^2(t,T))$ con ley normal conjunta

$$N_2(0, 0, \nu_{S_i}^2(t,T), \nu_V^2(t,T), \rho_i(t,T))^8.$$

Por la definición de esperanza condicional, se tiene

$$\begin{aligned}
E_1 &= S_1(t) \int_t^T \int_t^T \exp(\eta_{S_1} - \frac{1}{2}\nu_{S_1}^2) \mathbf{1}_{\left\{S_1^F(T,T) > S_2^F(T,T)\right\}} \mathbf{1}_{\{V_F(T,T) < D\}} d\eta_V d\eta_{S_1} \\
&= S_1(t) \int_t^T \int_t^T \mathbf{1}_{\left\{S_1^F(T,T) > S_2^F(T,T)\right\}} \mathbf{1}_{\{V_F(T,T) < D\}} d\dot{\eta}_V d\dot{\eta}_{S_1} \\
&= S_1(t) N_2(a_1, a_2, \rho_1(t,T))
\end{aligned}$$

donde $\dot{\eta}_i(t,T) = \int_t^T \beta_i(s,T) d\check{W}(s),\ \forall\ i \in \{S_1, S_2, V\}$, y a_1 y a_2 están determinadas por la evolución de las funciones indicadoras.

Por el Teorema de Girsanov, el cambio de medida por (14) implica que

$$\check{W}(t) = \tilde{W}(t) - \int_0^t \beta_{S_1}(s,T) ds$$

es un movimiento Browniano estándar bajo $\dot{\mathbf{F}}$. Las funciones indicadoras pueden por consiguiente ser evaluadas como

$$\begin{aligned}
&\dot{\mathbf{F}}\left(S_1^F(T,T) > S_2^F(T,T)\right) \\
&= \dot{\mathbf{F}}\left(S_1^F(t,T) \exp(\eta_{S_1}(t,T) - \frac{1}{2}\nu_{S_1}^2(t,T)) > S_2^F(t,T) \exp(\eta_{S_2}(t,T) - \frac{1}{2}\nu_{S_2}^2(t,T))\right) \\
&= \dot{\mathbf{F}}\left(\frac{S_1(t)}{P(t,T)} \exp\left(\int_t^T \beta_{S_1}(s,T) \cdot (d\check{W}(s) + \beta_{S_1}(s,T) ds) - \frac{1}{2}\nu_{S_1}^2(t,T)\right)\right. \\
&\qquad \left. > \frac{S_2(t)}{P(t,T)} \exp\left(\int_t^T \beta_{S_2}(s,T) \cdot (d\check{W}(s) + \beta_{S_1}(s,T) ds) - \frac{1}{2}\nu_{S_2}^2(t,T)\right)\right) \\
&= \dot{\mathbf{F}}\left(\ln S_1(t) + \eta_{S_1}(t,T) + \frac{1}{2}\nu_{S_1}^2(t,T)\right. \\
&\qquad \left. > \ln S_2(t) + \eta_{S_2}(t,T) + \int_t^T \beta_{S_1}(s,T) \cdot \beta_{S_2}(s,T) ds - \frac{1}{2}\nu_{S_2}^2(t,T)\right)
\end{aligned}$$

[8] Donde N_2 es la función de distribución acumulada de una variable aleatoria normal estándar conjunta bivariada definida por $N_2(x_1, x_2, \rho) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(z_1, z_2, \rho) dz_1 dz_2$, con densidad $f(z_1, z_2, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{2(1-\rho^2)}\right)$

$$
\begin{aligned}
&= \dot{\mathbf{F}}\left(\int_t^T (\beta_{S_1}(s,T) - \beta_{S_2}(s,T))d\check{W}(s)\right.\\
&\qquad \left. < \ln\frac{S_1(t)}{S_2(t)} + \frac{1}{2}\nu_{S_1}^2(t,T) - \frac{1}{2}\nu_{S_2}^2(t,T) - \int_t^T \beta_{S_1}(s,T)\cdot\beta_{S_2}(s,T)ds\right)\\
&= \dot{\mathbf{F}}\left(\zeta < \frac{\ln\frac{S_1(t)}{S_2(t)} + \frac{1}{2}\chi^2(t,T)}{\chi(t,T)}\right)
\end{aligned}
$$

con $\chi^2(t,T) = \int_t^T ||\beta_{S_1}(s,T) - \beta_{S_2}(s,T)||^2 ds = \nu_{S_1}^2(t,T) + \nu_{S_2}^2(t,T) - \phi_{S_1S_2}(t,T)$.

$$
\begin{aligned}
&\dot{\mathbf{F}}(V_F(t,T) > D)\\
&= \dot{\mathbf{F}}\left(V_F(t,T)\exp\left(\eta_V(t,T) - \frac{1}{2}\nu_V^2(t,T)\right) > D\right)\\
&= \dot{\mathbf{F}}\left(\frac{V(t)}{P(t,T)}\exp\left(\int_t^T \beta_V(s,T)\cdot(d\check{W}(s) + \beta_{S_1}(s,T)ds) - \frac{1}{2}\nu_V^2(t,T)\right) > D\right)\\
&= \dot{\mathbf{F}}\left(\eta_V(t,T) > \ln\frac{DP(t,T)}{V(t)} - \phi_{S_1V}(t,T) + \frac{1}{2}\nu_V^2(t,T)\right)\\
&= \dot{\mathbf{F}}\left(\eta_V(t,T) < \ln V(t) - \ln D - \ln P(t,T) + \phi_{S_1V}(t,T) - \frac{1}{2}\nu_V^2(t,T)\right)
\end{aligned}
$$

Por consiguiente,

$$E_1 = S_1(t)N_2(a_1, a_2, \rho_1(t,T)),$$

donde

$$
\begin{aligned}
a_1 &= \frac{\ln\frac{S_1(t)}{S_2(t)} + \frac{1}{2}\chi^2(t,T)}{\chi(t,T)}\\
a_2 &= \frac{\ln\frac{V(t)}{DP(t,T)} + \phi_{S_1V}(t,T) - \frac{1}{2}\nu_V^2(t,T)}{\nu_V(t,T)}
\end{aligned}
$$

Evaluación de E_2.

$$E_2 = P(t,T)I\!E_{\mathbf{F}}\left[S_2^F(t,T)\mathbf{1}_{\left\{S_1^F(t,T) > S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right]$$

La evolución de esta esperanza es similar a E_1 donde S_1 es reemplazado por S_2. En este caso, la medida auxiliar $\dot{\mathbf{F}}$ esta definida por consiguiente como

$$\zeta(t) = \frac{d\dot{\mathbf{F}}}{d\mathbf{F}} = \exp\left(\int_0^T \beta_{S_2}(s,T)d\check{W}(s) - \frac{1}{2}\int_0^T ||\beta_{S_2}(s,T)||^2 ds\right)$$

En consecuencia, la evolución de las funciones indicadoras cambian a

$$\begin{aligned}
&\dot{\mathbf{F}}\left(S_1^F(T,T) > S_2^F(t,T)\right)\\
&= \dot{\mathbf{F}}\left(\frac{S_1(t)}{P(t,T)}\exp\left(\int_t^T \beta_{S_1}(s,T)\cdot(d\check{W}(s)+\beta_{S_2}(s,T)ds)-\frac{1}{2}\nu_1^2(t,T)\right)\right.\\
&\quad \left. > \frac{S_2(t)}{P(t,T)}\exp\left(\int_t^T \beta_{S_2}(s,T)\cdot(d\check{W}(s)+\beta_{S_2}(s,T)ds)-\frac{1}{2}\nu_2^2(t,T)\right)\right)\\
&= \dot{\mathbf{F}}\left(\zeta < \frac{\ln\frac{S_1(t)}{S_2(t)}-\frac{1}{2}\chi^2(t,T)}{\chi(t,T)}\right)
\end{aligned}$$

La segunda función indicadora también esta evolucionando bajo la medida $\dot{\mathbf{F}}$, obteniéndose

$$\dot{\mathbf{F}}\left(V_F(t,T) > D\right) = \dot{\mathbf{F}}\left(\zeta < \frac{\ln\frac{V(t)}{DP(t,T)}+\phi_{S_2}(t,T)-\frac{1}{2}\nu_V^2(t,T)}{\nu_V}\right)$$

Por lo tanto, se sigue que

$$E_2 = S_2(t)N_2(b_1, b_2, \rho_2(t,T)),$$

donde

$$\begin{aligned}
b_1 &= \frac{\ln\frac{S_1(t)}{S_2(t)}-\frac{1}{2}\chi^2(t,T)}{\chi(t,T)}\\
b_2 &= \frac{\ln\frac{V(t)}{DP(t,T)}+\phi_{S_2}(t,T)-\frac{1}{2}\nu_V^2(t,T)}{\nu_V}.
\end{aligned}$$

Evaluación de E_3. Se evalúa la expresión

$$E_3 = P(t,T)I\!E_{\mathbf{F}}\left[S_1^F(T,T)V_F(T,T)D^{-1}\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)<D\}}|\mathcal{F}_t\right].$$

Sustituyéndose a $S_1^F(T,T)$ y $V_F(T,T)$, se obtiene

$$\begin{aligned}
E_3 &= I\!E_{\mathbf{F}}\left[\exp\left(\eta_{S_1}(t,T)-\frac{1}{2}\nu_{S_1}^2(t,T)\right)\exp\left(\eta_V(t,T)-\frac{1}{2}\nu_V^2(t,T)\right)\right.\\
&\quad \left.\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right]\frac{S_1(t)V(t)}{DP(t,T)}\\
&= \frac{S_1(t)V(t)}{DP(t,T)}I\!E_{\mathbf{F}}\left[\exp\left(\int_t^T(\beta_{S_1}(s,T)+\beta_V(s,T))d\check{W}(s)\right.\right.\\
&\quad \left.-\frac{1}{2}\int_t^T||\beta_{S_1}(s,T)||^2+||\beta_V(s,T)||^2ds\right)\\
&\quad \left.\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\geq D\}}|\mathcal{F}_t\right].
\end{aligned}\tag{15}$$

Ahora se introduce una nueva medida $\dot{\mathbf{F}}$ definida por

$$\frac{d\dot{\mathbf{F}}}{d\mathbf{F}} = \exp\left(\int_0^T \beta_{S_1}(s,T) + \beta_V(s,T)d\tilde{W}(s) - \frac{1}{2}\int_0^T ||\beta_{S_1}(s,T) + \beta_V(s,T)||^2 ds\right)$$

con $\dot{\zeta}(t) = I\!\!E_{\mathbf{F}}\left[\frac{d\dot{\mathbf{F}}}{d\mathbf{F}}|\mathcal{F}_t\right]\ \forall\ t \in [0,T]$. Puesto que

$$\int_t^T ||\beta_{S_1}(s,T) + \beta_V(s,T)||^2 ds$$

$$= \int_t^T ||\beta_{S_1}(s,T)||^2 + ||\beta_V(s,T)||^2 ds + 2\int_t^T \beta_{S_1}(s,T)\beta_V(s,T)ds$$

la expresión (15) puede ser igual a

$$\begin{aligned} E_3 &= \frac{S_1(t)V(t)}{DP(t,T)}\dot{\zeta}^{-1}(t)I\!\!E_{\mathbf{F}}\left[\dot{\zeta}(T)\exp(\phi_{S_1V}(t,T))\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\le D\}}|\mathcal{F}_t\right] \\ &= \frac{S_1(t)V(t)}{DP(t,T)}\exp(\phi_{S_1V}(t,T))\dot{\zeta}^{-1}(t)I\!\!E_{\mathbf{F}}\left[\dot{\zeta}(T)\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\le D\}}|\mathcal{F}_t\right] \end{aligned}$$

donde $\phi_{S_1V}(t,T) = \int_t^T \beta_{S_1}(s,T)\beta_V(s,T)ds$. Utilizando la regla de Bayes se tiene que

$$E_3 = \frac{S_1(t)V(t)}{DP(t,T)}\exp(\phi_{S_1V}(t,T))I\!\!E_{\dot{\mathbf{F}}}\left[\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(T,T)\right\}}\mathbf{1}_{\{V_F(T,T)\le D\}}|\mathcal{F}_t\right]$$

Por la definición de movimiento Browniano, $\eta_i(t,T)$ tiene una $\dot{\mathbf{F}}$ ley $N(0,\nu_i^2(t,T))$ con ley normal $N_2\left(0,0,\nu_{S_i}^2(t,T),\nu_V^2(t,T),-\rho_i(t,T)\right)$. Por consiguiente lo anterior se puede escribir

$$\begin{aligned} E_3 &= \frac{S_1(t)V(t)}{DP(t,T)}\int_t^T\int_t^T \exp\left(\eta_{S_1} - \frac{1}{2}\nu_{S_1}^2 + \eta_V - \frac{1}{2}\nu_V^2\right) \\ &\quad \mathbf{1}_{\left\{S_1^F(T,T)>S_2^F\right\}}\mathbf{1}_{\{V_F(T,T)<D\}}d\eta_V d\eta_{S_1} \\ &= \frac{S_1(t)V(t)}{DP(t,T)}\exp\left(\phi_{S_1V}(t,T)\right)\int_t^T\int_t^T \mathbf{1}_{\{S_F(T,T)>K\}}\mathbf{1}_{\{V_F(T,T)<D\}}d\dot{\eta}_V d\dot{\eta}_{S_1} \\ &= \frac{S_1(t)V(t)}{DP(t,T)}\exp\left(\phi_{S_1V}(t,T)\right)N_2(c_1,c_2,-\rho(t,T)) \end{aligned}$$

Las funciones indicadoras determinan a c_1 y c_2. Por el Teorema de Girsanov, el cambio de medida anterior implica que

$$\check{W}(t) = \tilde{W}(t) - \int_0^t \left(\beta_{S_1}(s,T) + \beta_V(s,T)\right)ds$$

es un movimiento Browniano estándar bajo $\dot{\mathbf{F}}$. Por lo tanto, las funciones indicadoras son evaluadas como sigue

$$
\begin{aligned}
&\dot{\mathbf{F}}\left(S_1^F(T,T) > S_2^F(T,T)\right)\\
=\;& \dot{\mathbf{F}}\left(S_1^F(t,T)\exp(\eta_{S_1}(t,T) - \frac{1}{2}\nu_{S_1}^2(t,T) > S_2^F(t,T)\exp(\eta_{S_2}(t,T) - \frac{1}{2}\nu_{S_2}^2(t,T)\right)\\
=\;& \dot{\mathbf{F}}\left(\frac{S_1(t)}{P(t,T)}\exp\left(\int_t^T \beta_{S_1}(s,T)\cdot(d\check{W}(s) + (\beta_{S_1}(s,T)+\beta_V(s,T))ds) - \frac{1}{2}\nu_{S_1}^2(t,T)\right)\right.\\
&\left. > \frac{S_2(t)}{P(t,T)}\exp\left(\int_t^T \beta_{S_2}(s,T)\cdot(d\check{W}(s) + (\beta_{S_1}(s,T)+\beta_V(s,T))ds) - \frac{1}{2}\nu_{S_2}^2(t,T)\right)\right)\\
=\;& \dot{\mathbf{F}}\left(\ln S_1(t) + \eta_{S_1}(t,T) + \frac{1}{2}\nu_{S_1}^2(t,T) + \phi_{S_1V}(t,T)\right.\\
&\left. > \ln S_2(t) + \eta_{S_2}(t,T) - \frac{1}{2}\nu_{S_2}^2(t,T) + \phi_{S_2V}(t,T) + +\phi_{S_1S_2}(t,T)\right)
\end{aligned}
$$

$$
\begin{aligned}
=\;& \dot{\mathbf{F}}\left(\int_t^T (\beta_{S_1}(s,T) - \beta_{S_2}(s,T))d\check{W}(s)\right.\\
&\left. < \ln\frac{S_1(t)}{S_2(t)} + \frac{1}{2}\nu_{S_1}^2(t,T) + \frac{1}{2}\nu_{S_2}^2(t,T) - \phi_{S_1S_2}(t,T) + \phi_{S_1V}(t,T) - \phi_{S_2V}(t,T)\right)\\
=\;& \dot{\mathbf{F}}\left(\zeta < \frac{\ln\frac{S_1(t)}{S_2(t)} + \frac{1}{2}\chi^2(t,T) + \phi_{S_1V}(t,T) - \phi_{S_2V}(t,T)}{\chi(t,T)}\right)
\end{aligned}
$$

con $\chi^2(t,T) = \int_t^T ||\beta_{S_1}(s,T) - \beta_{S_2}(s,T)||^2 ds$.

$$
\begin{aligned}
&\dot{\mathbf{F}}\left(V_F(t,T) > D\right)\\
=\;& \dot{\mathbf{F}}\left(V_F(t,T)\exp\left(\eta_V(t,T) - \frac{1}{1}\nu_V^2(t,T)\right) > D\right)\\
=\;& \dot{\mathbf{F}}\left(\frac{V(t)}{P(t,T)}\exp\left(\int_t^T \beta_V(s,T)\cdot(d\check{W}(s) + (\beta_{S_1}(s,T)+\beta_V(s,T))ds) - \frac{1}{2}\nu_V^2(t,T)\right) > D\right)\\
=\;& \dot{\mathbf{F}}\left(\frac{V(t)}{P(t,T)}\exp\left(\eta_V(t,T) + \int_t^T \beta_{S_1}(s,T)\cdot\beta_V(s,T)ds + \frac{1}{2}\nu_V^2(t,T)\right) > D\right)\\
=\;& \dot{\mathbf{F}}\left(\eta_V(t,T) > \ln\frac{DP(t,T)}{V(t)} - \phi_{S_1V}(t,T) - \frac{1}{2}\nu_V^2(t,T)\right)\\
=\;& \dot{\mathbf{F}}\left(\eta_V(t,T) < \ln V(t) - \ln D - \ln P(t,T) + \phi_{S_1V}(t,T) - \frac{1}{2}\nu_V^2(t,T)\right)
\end{aligned}
$$

Por lo tanto,

$$E_3 = S_1(t)N_2\left(c_1, c_2, -\rho_1(t,T)\right),$$

donde

$$
\begin{aligned}
c_1 &= \frac{\ln \frac{S_1(t)}{S_2(t)} + \frac{1}{2}\chi^2(t,T)}{\chi(t,T)} \\
c_2 &= \frac{\ln \frac{V(t)}{DP(t,T)} + \phi_{S_1V}(t,T) - \frac{1}{2}\nu_V^2(t,T)}{\nu_V(t,T)}
\end{aligned}
$$

Evaluación de E_4. Se evalúa la expresión

$$E_4 = P(t,T)I\!E_{\mathbf{F}}\left[S_2^F(T,T)V_F(T,T)D^{-1}\mathbf{1}_{\left\{S_1^F(T,T)>S_2^F(t,T)\right\}}\mathbf{1}_{\{V_F(T,T)<D\}}|\mathcal{F}_t\right].$$

La evaluación es similar a E_3, sin embargo, la nueva medida introducida esta definida por

$$\frac{d\dot{\mathbf{F}}}{d\mathbf{F}} = \exp\left(\int_0^T (\beta_{S_2}(s,T)+\beta_V(s,T))\tilde{W}(s) - \frac{1}{2}\int_0^T ||\beta_{S_2}(s,T)+\beta_V(s,T)||^2 ds\right).$$

La evaluación de la función indicadora en consecuencia cambia, por lo tanto, da como resultado

$$E_4 = S_2(t)N_2\left(d_1, d_2, -\rho_2(t,T)\right),$$

donde

$$
\begin{aligned}
d_1 &= \frac{\ln \frac{S_1(t)}{S_2(t)} - \frac{1}{2}\chi^2(t,T)}{\chi(t,T)} \\
d_2 &= \frac{\ln \frac{V(t)}{DP(t,T)} + \phi_{S_1V}(t,T) - \frac{1}{2}\nu_V^2(t,T)}{\nu_V(t,T)}
\end{aligned}
$$

$\square$

Proceso de Markov

En este capítulo se dan las nociones de lo que es una cadana de Markov a tiempo continuo, su estructura, la matriz infinitesimal, entre otras cosa. En las secciones subsecuentes se presenta una breve introducción al concepto de calificaciones crediticias en la cual se describen los principales resultados sobre matrices de transición y se examinan diferentes métodos para estimar y comparar matrices de transición crediticias.

Procesos en tiempo continuo

Un proceso estocástico es una colección de variables aleatorias definidas en un mismo espacio de probabilidad con valores en un conjunto arbitraruio. Esta colección va a depender de otro conjunto de subíndices llamado conjunto de parámetros, donde dicho conjunto va a tener cierto contenido dependiendo de las circunstancias del problema.

Definición 1 *Un proceso estocástico con espacio de estados E es una colección* $\{S_t : t \in T\}$ *de variables aleatorias* S_t *definidas en un espacio de probabilidad y que toman valores en E. El conjunto* T [1] *es llamado el conjunto de parámetros. Si* T *es numerable, en particular si* $T = \mathbb{N}$, *el proceso es llamado* ***proceso con parámetro discreto****, en cambio si* T *es no numerable, en particular* $T = \mathbb{R}_+$, *entonces el proceso es llamado* ***proceso con parámetro continuo****.*

Existe un tipo especial de procesos estocásticos con la propiedad de que el futuro es independiente del pasado dado el presente; a estos procesos se les denominan *Procesos de Markov.*

Definición 2 *Un proceso estocástico* $\{X_t : t \in \mathbb{R}_+\}$ *es un Proceso de Markov con espacio de estados* E *si para toda* $t, s \geq 0$ *y* $j \in E$

$$\mathbb{P}\left(X_{t+s} = j | X_u; u < t\right) = \mathbb{P}\left(X_{t+s} = j | X_t\right)$$

La ecuacún anterior puede en muchas ocaciones depender tanto de t como de s, pero hay veces en donde la probabilidad no depende de donde se este ubicado, es decir,

$$\mathbb{P}\left(X_{t+s} = j | X_t = i\right) = p_{i,j}(s)$$

cuando esto sucede, el proceso es llamado *Proceso de Markov de tiempo homogéneo.*

[1] Es usual interpretar al conjunto T como el conjunto del tiempo.

Cadenas de Markov en tiempo continuo

Una cadena de Markov en tiempo continuo $\{X_t : t \in \mathbb{R}_+\}$ es un proceso de Markov en el espacio de estados finito E.

La propiedad Markoviana afirma que $p_{i,j}(t)$ satisface

i) $p_{ij}(t) \geq 0 \quad t > 0$,

ii) $\sum_j p_{ij}(t) = 1 \quad t > 0$,

iii) $\sum_k p_{ik}(t)p_{kj}(s) = p_{ij}(t+s) \quad t, s > 0$,

iv) $\lim_{t\to 0} p_{ij}(t) = \begin{cases} 1 & si \quad i = j \\ 0 & si \quad i \neq j \end{cases}$

Si $\mathbf{P}(t) \overset{def}{=} \{p_{i,j}(t)\}_{i,j\in E}$ denota a la matriz de transición, entonces la familia $\{\mathbf{P}(t)\}_{t\geq 0}$ es el semigrupo de transición de la cadena de Markov homogénea en tiempo continuo.

Si el proceso $\{X(t)\}_{t\geq 0}$ es una cadena de Markov homogénea en tiempo continuo entonces, su semigrupo de transición es

$$\mathbf{P}(t) = \sum_{n=0}^{N} \exp(-\lambda t)\frac{(\lambda t)^n}{n!}\mathbf{K}^n,$$

esto es,

$$p_{ij}(t) = \sum_{n=0}^{N} \exp(-\lambda t)\frac{(\lambda t)^n}{n!}k_{ij}(n).$$

Matriz infinitesimal

Sea $\{\mathbf{P}(t)\}_{t\geq 0}$ un semigrupo de transición en E, i.e., para cada $t, s \geq 0$,

(a) $\mathbf{P}$(t) es una matriz estocástica,

(b) $\mathbf{P}(0) = \mathrm{I}$,

(c) $\mathbf{P}(\mathrm{t+s}) = \mathbf{P}(\mathrm{t})\mathbf{P}(\mathrm{s})$

Supóngase, aún más, que el semigrupo es *continuo* en el origen, es decir,

$$\lim_{h\downarrow 0} \mathbf{P}(h) = \mathbf{P}(0) = I,$$

esto implica continuidad en cualquier tiempo $t \geq 0$, es decir,

$$\lim_{h\to 0} p_{ij}(t+h) = p_{ij}(t)$$

para todo estado i, j.

Teorema 1 *Sea $\{\boldsymbol{P}(t)\}_{t\geq 0}$ un semigrupo de transición continuo en el espacio de estados E. Para algún estado i, existe*

$$-\lambda_{ii} \stackrel{def}{=} \lim_{t\to 0} \frac{1-p_{ii}(t)}{t} \stackrel{def}{=} \lambda_i,$$

y para algún par de i, j de diferente estado, existe

$$\lambda_{ij} \stackrel{def}{=} \lim_{t\to 0} \frac{p_{ij}(t)}{t},$$

La demostración se puede consultar en el libro de Brémaud.[**?**]

Definición 3 *Los números λ_{ij} son llamados característicos locales del semigrupo, o de la correpondiente cadena de Markov homogénea en tiempo continuo. La matriz*

$$\boldsymbol{\Lambda} = \{\lambda_{\mathbf{ij}}\}_{\mathbf{i,j\in E}}$$

es llamada la matriz generadora infinitesimal de un semigrupo, o de una cadena de Markov homogénea en tiempo continuo.

En notación compacta

$$\boldsymbol{\Lambda} = \lim_{\mathbf{h}\downarrow\mathbf{0}} \frac{\mathbf{P(h)}-\mathbf{P(0)}}{\mathbf{h}}$$

En este sentido la matriz generadora infinitesimal $\boldsymbol{\Lambda}$ es la derivada en cero de la función matriz $t \mapsto \mathbf{P}(t)$.

Definición 4 *Si para todo estado i,*

$$\lambda_i < \infty \tag{1}$$

el semigrupo $\{\boldsymbol{P}(t)\}$ es llamado estable. Si para todo estado i,

$$\lambda_i = \sum_{\substack{j\in E\\ j\neq i}} \lambda_{ij} \tag{2}$$

este es llamado conservador.

La razón por la última denominación viene de la conservación de la igualdad

$$\sum_{j\in E} p_{ij}(h) = 1,$$

o equivalentemente,

$$\frac{1-p_{ii}(h)}{h} = \sum_{\substack{j\in E\\ j\neq i}} \frac{p_{ij}(h)}{h},$$

del cual se deduce

$$\lambda_{ij} = \lim_{h\downarrow 0} \sum_{\substack{j\in E\\ j\neq i}} \frac{p_{ij}(h)}{h}.$$

El intercambio de la suma y el límite es siempre posible si E es finito. En este caso (2) vale, y consecuentemente, (1) es valido porque λ_{ij} es finito para todo par de estados i, j. Además, el semigrupo de una cadena de Markov homogénea uniforme es estable y conservadora, incluso cuando E es infinito.

Sistemas diferenciales Kolmogorov

En vista de las propiedades de los semigrupos, para toda $t \geq 0$ y toda $h \geq 0$

$$\frac{\mathbf{P}(t+h) - \mathbf{P}(t)}{h} = \mathbf{P}(t)\frac{\mathbf{P}(h) - I}{h} = \frac{\mathbf{P}(h) - I}{h}\mathbf{P}(t) \tag{3}$$

Como el espacio de estados E es finito, entonces se puede tomar el límite en (3), obteniéndose el sistema diferencial

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{\Lambda} = \mathbf{\Lambda}\mathbf{P}(\mathbf{t}),$$

donde $\mathbf{\Lambda}$ es la matriz generadora infinitesimal. La ecuación

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{\Lambda}\mathbf{P}(\mathbf{t}) \tag{4}$$

puede ser escrita explícitamente. Para toda $i, j \in E$,

$$\frac{d}{dt}p_{ij}(t) = -\lambda_{ij}p_{ij}(t) + \sum_{\substack{k\in E\\ k\neq i}} \lambda_{ik}p_{kj}(t).$$

El sistema (4) es conocido como el sistema diferencial *backward* de Kolmogorov. El sistema diferencial *forward* es

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\mathbf{\Lambda}. \tag{5}$$

esto es, para toda $i, j \in E$,

$$\frac{d}{dt}p_{ij}(t) = -p_{ij}(t)\lambda_{ij} + \sum_{\substack{k\in E\\ k\neq i}} p_{ik}(t)\lambda_{kj}.$$

Donde el espacio de estados es finito, la solución de (4) o (5) con la condición inicial $\mathbf{P}(0) = I$ es

$$\mathbf{P}(t) = e^{t\mathbf{\Lambda}} \tag{6}$$

donde la exponencial de una matriz de dimensión finita $\mathbf{C}$ esta definida por

$$e^{\mathbf{C}} \stackrel{def}{=} \sum_{n=0}^{\infty} \frac{\mathbf{C}^{\mathbf{n}}}{n!}$$

Propiedad fuerte de Markov

Una cadena de Markov homogénea tiene la propiedad fuerte de Markov. Para una meticulosa afirmación de este resultado, es necesario introducir el concepto de *tiempo de espera* en tiempo continuo.

Definición 5 *Sea $\{X(t)\}_{t\geq 0}$ un proceso estocástico con valores en E. Una variable aleatoria τ tomando sus valores en $\mathbb{R}_+$ es llamada tiempo de espera con respecto a $\{X(t)\}$ sí para toda $t \geq 0$, el $\{\tau \leq t\}$ es expresado en términos de $(X(s),\ s \in [0,t])$. Se denotará esto por*

$$\{\tau \leq t\} \in X_0^t$$

Teorema 2 *Sea $\{X(t)\}_{t\geq 0}$ una cadena de Markov homogénea en tiempo continuo con espacio de estado $E = \mathbb{N}$ (sin perdida de generalidad) y semigrupo de transición $\{\boldsymbol{P}(t)\}_{t\geq 0}$, y sea τ un tiempo de espera con respecto a $\{X(t)\}_{t\geq 0}$. Sea $k \in E$ un estado arbitrario. Entonces, dado que $X(\tau) = k$,*

a) *La cadena despúes de τ y la cadena antes de τ son independientes, y*

b) *la cadena después de τ es una cadena de Markov homogénea en tiempo continuo de salto regular con semigrupo de transición $\{\boldsymbol{P}(t)\}_{t\geq 0}$.*

La demostración puede consultarse en el libro de Brémaud. [**?**]

Calificadoras y calificaciones

Las calificaciones se deben interpretar como opiniones de las agencias calificadoras como *Standard and Poor's (S&P)* y *Moody's Investor Services (Moody's)* de la calidad crediticia de un deudor con respecto al cumplimiento de sus obligaciones o compromisos en especifico. Una calificación de crédito representa un tasa global del cumplimiento de las obligaciones por parte del acreditado.

Existen algunas diferencias entre las calificaciones de las diferentes compañías. Desde el punto de vista de la administración de riesgo de crédito se dice que S&P mide la probabilidad de incumplimiento, mientras Moody's mide la distancia a la pérdida o a la esperanza de pérdida.

Las calificaciones pueden aplicarse al emisor y a alguna emisión. En el emisor lo que se evalúa es la capacidad y voluntad para cumplir con sus compromisos financieros, basándose principalmente en el riesgo de incumplimiento, y que pueder ser evaluado de la siguiente forma

$$\text{Riesgo de Incumplimiento} \begin{cases} \text{Calificaciones de riesgo Corporativo} \\ \text{Calificaciones de riesgo de la Contraparte} \\ \text{Calificaciones de riesgo Soberano} \end{cases}$$

Por otra parte cuando alguna agencia califica una emisión, ésta evalúa la calidad crediticia de una obligación financiera específica, sin dejar a un lado la calidad crediticia de su emisor y adicionalmente se toman en particular algunos factores que influyen en la decisión de otorgar una calificación a la emisión en cuestión.

Existen varias escalas de calificaiones entre las agencias más reconocidad, las calificaciones de crédito son indicadas por símbolos, cada símbolo representa un grupo en que las características del crédito son ampliamente las mismas. Hay nueve símbolos como los que se muestran a continuación:

Aaa Aa A Baa Ba B Caa Ca C

Algunas calificadoras añaden modificadores numéricos 1, 2, y 3 a cada clasificación genérica de los ratings desde Aaa hasta Caa. El modificador 1 indica que la emisión esta en un rango alto de su categoría; el modificador 2 indica que se encuentra en un rango medio; y el modificador 3 un rango bajo de esa categoría.

La calidad del crédito de la mayoría de los emisores y sus obligaciones no es fija durante un período de tiempo, sino que tiende a sufrir cambios. Por esta razón cambian en las calificaciones para reflejar las variaciones en la posición relativa intrínseca de los emisores y sus obligaciones. Un cambio en la calificación puede ocurrir cuando quiera en el caso de emisiones individuales. Tal calificación de cambio debe servir como aviso de que se observa alguna alteración en la credibilidad, o que la calificación anterior no reflejó totalmente la calidad de la emsión. Mientras debido a su misma naturaleza, los cambios serán esperados frecuentemente más entre las emisiones de más baja calificaión que entre las emisiones de calificaciones más altas.

Estas calificaciones que otorgan las compañías calificadoras en general son:

- La opinión de incurrir en incumplimiento.
- Una predicción del grado relativo de cobertura que un inversionista tiene si un acreditado sufre una contariedad económica
- Una medida de probabilidad de que un acreditado pueda incidir en incumplimiento con respecto a sus obligaciones en el contrato de crédito durante la duración del mismo.
- Una medida comparable, de riesgo de incumplimiento entre un amplio rango de posibles acreditados.

Matrices de transición de crédito

Los modelos de riesgo de crédito típicamente asumen que, a través del tiempo, las calificaciones de crédito emigran entre un conjunto de posibles estados

[2] Fuente: Moody's Investor Services

Calificación	Categoría de clasificación de deuda a largo plazo [2]
Aaa	Muy alta capacidad de cumplimiento, la cual no se vería afectada ante posibles cambios en el emisor, en el sector al que pertenece o en la eonomía.
Aa	Alta capacidad de cumplimiento, la cual no se vería afectada ante posibles cambios en el emisor, en el sector al que pertenece o en la eonomía.
A	Buena capacidad de cumplimiento, la cual es susceptible a deteriorarse levemente ante posibles cambios en el emisor, en el sector al que pertenece o en la eonomía.
Baa	Suficiente capacidad de cumplimiento, la cual es susceptible a debilitarse ante posibles cambios en el emisor, en el sector al que pertenece o en la eonomía.
Ba	Cuenta con capacidad de cumplimiento, la cual es variable ante posibles cambios en el emisor, en el sector al que pertenece o en la eonomía, pudiendo incurrir en retraso de pago.
B	Cuenta con una mínima capacidad de cumplimiento, siendo muy variable y susceptible ante posibles cambios en el emisor, en el sector al que pertenece o en la eonomía, pudiendo incurrir en incumplimiento.
Caa	No cuenta con la capacidad de cumplimiento, existiendo alto riesgo de pérdida.
Ca	No cuenta con la capacidad de cumplimiento, y representa incumplimiento efectivo.
C	Carencia de garantía suficiente, con poca perspectiva de recuperar la inversión.

de crédito. Matemáticamente, este proceso puede ser modelado como una cadena de Markov finita, la cuál asume que las calificaiones de crédito cambian sucesivamente de un estado a otro, en un intervalo de tiempo dado, con una cierta probabilidad. Las probabilidades de migración de crédito de una matriz de transición. Las matrices de transición, las cuales espefician la probabilidad de cambio de estado en un periodo de tiempo, pueden ser obtenidas de varias agencias especializadas en datos de crédito.

Una aproximación que obtiene matrices de transición por periodos de longitud arbitraria involucra la inclusión de cadenas de Markov discretas en un proceso de Markov continuo. Para un proceso de Markov continuo, cualquier matriz de transición puede ser expresada como la exponencial de una matriz generadora. Así, resolviendo el *problema de inclusión* esencialmente permite encontrar una generadora que es consistente con la matriz de transición de una cadena de Markov discreta.

Problema de inclusión

Dada una cadena de Markov $\{X_t : t \in \mathbb{N}\}$ con espacio de estados finito, el problema de inclusión puede ser planteado como:

> Es posible construir un proceso de Markov $\left\{\tilde{X}_\tau : \tau \in \mathbb{R}\right\}$ tal que la distribución de probabilidad de $\tilde{X}_\tau$ al tiempo $\tau \in \mathbb{N}$ es igual a la distribución de X_τ.

La principal idea desarrollada es basada en la noción de la *generadora* de un proceso de Markov, por generadora se deberá entender a una matrix $\mathbf{\Lambda}$ cuyos elemetos satisfacen

$$\begin{aligned} \sum_{j=i}^{n} \lambda_{ij} = 0 \qquad \text{para} \;\; i \in E \\ \lambda_{ij} \geq 0 \qquad \text{para} \;\; i,j \in E \;\; y \;\; i \neq j \end{aligned} \tag{7}$$

Sea $\tilde{\mathbf{P}}(\tau)$ una matriz de transición del proceso de Markov $\tilde{X}(\tau)$. La matriz $\tilde{\mathbf{\Lambda}} = (\lambda_{ij})_{i,j\in E}$ es llamada la matriz generadora de $\tilde{X}(\tau)$ si $\tilde{\mathbf{P}}(\tau)$ satisface

$$\frac{d}{d\tau}\tilde{\mathbf{P}} = \tilde{\mathbf{P}}\tilde{\mathbf{\Lambda}}$$

En este caso, se obtiene

$$\tilde{\mathbf{P}}(\tau) = e^{\tau\tilde{\mathbf{\Lambda}}}, \quad \tau > 0$$

Por consiguiente, si la distribución del proceso de Markov $\tilde{X}(\tau)$ es idéntica a la de la cadena de Markov X_t al tiempo $\tau = 1, 2, \ldots$ entonces la matriz de transición $\mathbf{P}$ debe satisfacer

$$\mathbf{P} = e^{\tilde{\mathbf{\Lambda}}}$$

Si la matriz $\tilde{\mathbf{\Lambda}}$ existe y es única, entonces se pude definir una extensión al tiempo continuo, $\tilde{\mathbf{P}}(\tau)$, de la matriz de transición $\mathbf{P}$ como sigue

$$\tilde{\mathbf{P}}(\tau) = e^{\tau\tilde{\mathbf{\Lambda}}}$$

Nótese que, puesto que τ es continua, entonces de la relación anterior se obtiene una matriz de transición para un periodo de longitud arbitrario.

Mientras se calcula $\tilde{\mathbf{\Lambda}} = \ln(\mathbf{P})$, no se garantiza que la matriz $\ln(\mathbf{P})$ satiface la ecuación (7). En general, la cuestion de cuando una matriz de transición $\mathbf{P}$ puede ser representada in la forma $\mathbf{P} = e^{\mathbf{\Lambda}}$, donde $\mathbf{\Lambda}$ es una matriz generadora, es técnicamente muy difícil.

Problema de la no existencia de una generadora

Es natural preguntarse en esta situación, si es posible determinar la no existencia de una generadora para una matriz de transición $\mathbf{P}$ dada bajo ciertas condiciones, afortunadamente una respuesta a esta pregunta puede ser encontrada en el siguiente teorema:

Teorema 3 *Sea $\boldsymbol{P}$ una matriz de transición, y sugóngase que*

a) *$det(\boldsymbol{P}) \leq 0$, o*

b) *$det(\boldsymbol{P}) > \prod_i p_{ii}$, o*

c) *hay estados i y j tal que j es accesible [3] desde i, pero $p_{ij} = 0$.*

entonces no existe una matriz generadora exacta para $\boldsymbol{P}$.

Demostración

La demostración puede consultarse en [**?**]

Problema de la unicidad de una generadora

Junto al problema de la existencia de la matrix generadora, existe también el problema de su unicidad, puesto que $\mathbf{P}$ a veces tiene más de una matriz generadora valida, como lo demuestra la siguiente proposición.

Proposición 1 *Existen matrices de transición $\boldsymbol{P}$ las cuales tienen más de una matriz generadora valida.*

Demostración

Sea

$$\mathbf{P} = \begin{pmatrix} \frac{2+3b}{5} & \frac{2-2b}{5} & \frac{1-b}{5} \\ \frac{2-2b}{5} & \frac{2+3b}{5} & \frac{1-b}{5} \\ \frac{2-2b}{5} & \frac{2-2b}{5} & \frac{1+4b}{5} \end{pmatrix}$$

donde $b = e^{-4\pi}$, tal que

$$\mathbf{\Lambda_1} = \begin{pmatrix} -2\pi & 2\pi & 0 \\ 0 & -2\pi & 2\pi \\ 4\pi & 0 & -4\pi \end{pmatrix}$$

$$\mathbf{\Lambda_2} = \begin{pmatrix} \frac{-12\pi}{5} & \frac{8\pi}{5} & \frac{4\pi}{5} \\ \frac{8\pi}{5} & \frac{-12\pi}{5} & \frac{4\pi}{5} \\ \frac{8\pi}{5} & \frac{8\pi}{5} & \frac{-16\pi}{5} \end{pmatrix}$$

[3] Un estado j es accesible desde un estado i si para algún entero $n \geq 0$, $p_{ij}^n > 0$.

puede ser verificado que $\mathbf{\Lambda_1}$ y $\mathbf{\Lambda_2}$ son generadoras, y $\exp(\mathbf{\Lambda_1}) = \exp(\mathbf{\Lambda_2}) = \mathbf{P}$. Así, $\mathbf{\Lambda_1}$ y $\mathbf{\Lambda_2}$ son ambas matrices generadoras validas de $\mathbf{P}$. $\square$

El resultado anterior origina la siguiente pregunta: ¿Cuando una matriz de transición $\mathbf{P}$ tiene múltiples matrices generadoras validas $\mathbf{\Lambda}$, cuál generadora es la "mejor"? [4] Para poder esclarecer un poco este problema, se puede observar empíricamente que es generalmente improbable para una calificación migrar a una calificación remota en un periodo corto de tiempo. Así, un método para escoger entre matrices generadoras validas es tomar la que tenga el valor más pequeño de

$$J = \sum_{i,j} |j-i||\lambda_{ij}|$$

A pesar de las dificultades originadas en la proposición 1, es algunas veces posible demostrar la unicidad de generadoras para $\mathbf{P}$, tal como los siguientes resultados lo demuestran.

Teorema 4 *Sea $\boldsymbol{P}$ una matriz de transición*

a) *Si $det(\boldsymbol{P}) > 1/2$, entonces $\boldsymbol{P}$ tiene a lo más una generadora.*

b) *Si $det(\boldsymbol{P}) > 1/2$ y $||\boldsymbol{P}-\boldsymbol{I}|| < 1/2$, entonces la única posible matriz generadora para $\boldsymbol{P}$ es $\ln(\boldsymbol{P})$.*

c) *Si $\boldsymbol{P}$ tiene distintos eigenvalores y $det(\boldsymbol{P}) > e^{-\pi}$, entonces la única posible matriz generadora para $\boldsymbol{P}$ es $\ln(\boldsymbol{P})$.*

Demostración

La demostración puede consultarse en [?]

Teorema 5 *Sea $\boldsymbol{P}$ una matriz de transición que tiene eigenvalores reales distintos.*

- *Si todos los eigenvalores de $\boldsymbol{P}$ son positivos, entonces $\ln(\boldsymbol{P})$ es la única matriz real $\mathbf{\Lambda}$ tal que $\exp(\mathbf{\Lambda}) = \boldsymbol{P}$.*
- *Si $\boldsymbol{P}$ tiene algún eigenvalor negativo, entonces no existe matriz real $\mathbf{\Lambda}$ tal que $\exp(\mathbf{\Lambda}) = \boldsymbol{P}$.*

Demostración

La demostración puede consultarse en [?]

[4] Esta pregunta es importante, debido a que diferentes generadoras conducen a diferentes valores de $\mathbf{P}_t = \exp(t\mathbf{\Lambda})$ para t suficientemente pequeña.

1. Métodos de estimación para matrices generadoras

Se han construido modelos que consideran las diferentes clases del crédito caracterizado por sus calificaciones y que permite moverse dentro de estas clases. En esta sección, se describirá la idea básica de estos modelos.

A continuación se examinarán algunos métodos de estimación para matrices generadoras de una matriz de transición.

Método JLT

Jarrow, Lando y Turnbull (1997) proponen un modelo que relaciona la probabilidad de incumplimiento con las calificaciones de créditos. Usan cadenas de Markov con espacios de estados finitos de tiempo homógeneos con una matriz generadora, además usan la medida de martingala $\mathbf{Q}$ [5].

Este modelo es una extensión del trabajo de Jarrow-Turnbull (1995), caracterizando el proceso de bancarrota como una cadena de Markov con espacio de estados finito $E = 1, 2, \ldots, K$. El espacio de estados E representa las diferentes clases de calificaciones. Mientras 1 es para la mejor calificación de crédito, K representa el caso de incumplimiento, además elimina el supuesto de que la intensidad de incumplimiento es constante todo el tiempo, aunque el supuesto de independencia y que la tasa de recuperación es constante y exógena, se mantienen. Así, la cadena de Markov homogénea con espacio de estados finito $\{X_t : 0 \leq t \leq \tau\}$ es representada por una matriz de transición $K \times K$:

$$\mathbf{P} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,K} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,K} \\ \cdots & \cdots & \cdots & \cdots \\ p_{K-1,1} & p_{K-1,2} & \cdots & p_{K-1,K} \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

donde $p_{i,j} \geq 0$ para toda $i, j,\ i \neq j$, y $p_{i,i} \equiv 1 - \sum_{j \neq i} p_{i,j}$ para toda $i, j \in E$. $p_{i,j}$ representa la probabilidad de ir desde el estado i al estado j en un paso.

Un problema en el modelo es que se supone que las compañías con la misma calificación tienen el mismo rendimiento en la prima de riesgo. Además, la prima de riesgo de crédito sólo se altera cuando hay transición. Lo anterior no se refleja en el precio de mercado, pero, afecta más tarde a la calificación.

Este modelo es útil para valuar y cubrir deudas coorporativas, derivados OTC con contraparte riesgosa, bonos guvernamentales sujetos a riesgo de incumplimineto, crédito de derivados y para administración de riesgo.

En el modelo JLT se hacen los siguientes supuestos fundamentales.

[5] Consultar apéndice

- Las tasas de interés y el proceso de incumplimiento son independientes bajo la medida de martingala $\mathbf{Q}$.
- La matriz de transición es homogénea, es decir, las probabilidades de transición no dependen de t, esto es, $p_{i,j}(t,t+1) = p_{i,j}(0,1) = p_{i,j}$.
- La matriz de transición tiene un estado especial que representa el evento de incumplimiento. Este estado es absorbente; una vez alcanzado, se permanecerá ahí indefinidamente. Entonces se tiene $p_{K,K} = 1$ y $p_{K,j} = 0$ para toda $j \neq K$.

Bajo la medida de martingala la matriz de transición de un periodo equivale a:

$$\tilde{\mathbf{P}}_{t,t+1} = \begin{pmatrix} \tilde{p}_{1,1}(t,t+1) & \tilde{p}_{1,2}(t,t+1) & \cdots & \tilde{p}_{1,K}(t,t+1) \\ \tilde{p}_{2,1}(t,t+1) & \tilde{p}_{2,2}(t,t+1) & \cdots & \tilde{p}_{2,K}(t,t+1) \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{p}_{K-1,1}(t,t+1) & \tilde{p}_{K-1,2}(t,t+1) & \cdots & \tilde{p}_{K-1,K}(t,t+1) \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

donde $\tilde{p}_{i,j}(t,t+1) \geq 0 \ \forall \ i,j, \ i \neq j$, y $\tilde{p}_{i,j}(t,t+1) \equiv 1 - \sum_{j \neq i} \tilde{p}_{i,j}(t,t+1)$ para toda $i,j \in E$. $\tilde{p}_{i,j}(t,t+1) > 0$ si y solo si $p_{ij} > 0$ para $0 \leq t \leq \tau - 1$. Sin restricciones adicionales, las probabilidades $\tilde{p}(t,t+1)$ pueden depender del pasado no inmediato, es decir, la propiedad de Markov no es satisfecha más. Por consiguiente, se asume que los ajustes a las primas de riesgo son tal que el proceso de calificaión bajo las probabiliades martingala satisfacen

$$\tilde{p}_{ij}(t,t+1) = \pi_i(t) p_{ij} \qquad \text{para toda } i,j \ i \neq j$$

donde $\pi_i(t)$ es una función determinista del tiempo tal que $\tilde{p}_{ij}(t,t+1) \geq 0$ para toda $i,j \ i \neq j$ y $\sum_{j \neq i} \tilde{p}_{ij}(t,t+1) \leq 1$ para $i \in E$.

El método JLT para estimar la matriz generadora asume que cada empresa hizo a lo más una transición a lo largo del año. Bajo esta hipótesis se puede mostrar que para $\lambda_{i,i} \neq 0$ con $i = 1, \ldots, K-1$

$$\exp(\mathbf{\Lambda}) = \begin{pmatrix} e^{\lambda_{1,1}} & \frac{\lambda_{1,2}\left(e^{\lambda_{1,1}}-1\right)}{\lambda_{1,1}} & \cdots & \frac{\lambda_{1,K}\left(e^{\lambda_{1,1}}-1\right)}{\lambda_{1,1}} \\ \frac{\lambda_{2,1}\left(e^{\lambda_{2,2}}-1\right)}{\lambda_{2,2}} & e^{\lambda_{2,2}} & \cdots & \frac{\lambda_{2,K}\left(e^{\lambda_{2,2}}-1\right)}{\lambda_{2,2}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\lambda_{K-1,1}\left(e^{\lambda_{K-1,K-1}}-1\right)}{\lambda_{K-1,K-1}} & \frac{\lambda_{K-1,2}\left(e^{\lambda_{K-1,K-1}}-1\right)}{\lambda_{K-1,K-1}} & \cdots & \frac{\lambda_{K-1,K}\left(e^{\lambda_{K-1,K-1}}-1\right)}{\lambda_{K-1,K-1}} \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

la estimación de $\hat{\mathbf{\Lambda}}$ puede ser obtenida resolviendo el siguiente sistema:

$$p_{i,i} = \exp \hat{\lambda}_{i,i} \qquad \text{para } i = 1, \cdots, K-1$$

$$p_{i,j} = \hat{\lambda}_{i,j}\left(\exp \hat{\lambda}_{i,i} - 1\right) \qquad \text{para } i,j = 1, \cdots, K-1$$

El modelo JLT proporciona la solución a este sistema como

$$\hat{\lambda}_{i,i} = \ln p_{i,i} \qquad \text{para } i = 1, \cdots, K-1$$

$$\hat{\lambda}_{i,j} = p_{i,j} \frac{\ln p_{i,i}}{p_{i,i} - 1} \qquad \text{para } i, j = 1, \cdots, K-1$$

Esto sólo conduce a una matriz generadora aproximada, sin embargo se esta garantizando que la generadora no tendrá entradas negativas fuera de la diagonal.

Por construcción, el estado 1 es visto como la mejor calificación crediticia y $K-1$ es la peor calificación crediticia antes de el estado de incumplimiento. Para asegurarse que la cadena de Markov usada para el modelo refleja el hecho de que "las más bajas calficaciones crediticias son las más arriesgadas", existe una condición simple en la matriz generadora que se puede checar.

Lema 1 *Sea* $\mathbf{\Lambda}$ *una matriz generadora valida de* $\boldsymbol{P}$*. Entonces las siguientes dos condiciones son equivalentes:*

a) $\sum_{j \leq k} p_{ij}$ *es una función no decreciente de* i *para cualquier* k *fija.*

b) $\sum_{j \leq k} \lambda_{ij} \leq \sum_{j \leq k} \lambda_{i+1,j}$ *para toda* i *y* k *tal que* $k \neq i+1$*.*

Demostración

La demostración puede consultarse en [**?**].

Método IRW

Dada la matriz de transición $\mathbf{P}$ de $K \times K$ se esta interesado en encontrar una matriz generadora $\mathbf{\Lambda}$ tal que se cumpla (5) con $t = 1$ sin perdida de generalidad. Tratando con la cuestión de si existe una matriz generadora para una matriz de transición dada $\mathbf{P}$, el punto de partida de la estimación se da por el siguiente teorema. Previo a esto se requiere calcular a S con

$$S = \text{máx}\left\{(a-1)^2 + b^2;\ \text{donde } a + bi \text{ es un eigenvalor de } \mathbf{P},\ a, b \in \mathbb{R}\right\}$$

donde todos los eigenvalores de $\mathbf{P}$ son examinados calculando el cuadrado del valor absoluto de los eigenvalores menos 1 y tomando el máximo de éstos.

Teorema 6 *Sea* $\boldsymbol{P}$ *una matriz de transición de* $K \times K$*, y supóngase que* $S < 1$ [6]*, entonces la serie*

$$\tilde{\mathbf{\Lambda}} = (\boldsymbol{P} - \boldsymbol{I}) - \frac{(\boldsymbol{P} - \boldsymbol{I})^2}{2} + \frac{(\boldsymbol{P} - \boldsymbol{I})^3}{3} - \frac{(\boldsymbol{P} - \boldsymbol{I})^4}{4} + \cdots \tag{8}$$

converge geométricamente rápido, y da origen a una matriz $\tilde{\mathbf{\Lambda}}$ *de* $K \times K$ *tal que (7) y satisfaciendo exactamente que* $\boldsymbol{P} = \exp(\tilde{\mathbf{\Lambda}})$*.*

[6] Se ha observado que no es tan importante checar la condición $S < 1$. De hecho, con tal que la serie (8) *converja absolutamente*, $\tilde{\mathbf{\Lambda}}$ cumple con (7) y (5) automáticamente, así que esta condición no esta necesaria.

Demostración

La demostración puede consultarse en [?]

Es importante notar que incluso si $\tilde{\mathbf{\Lambda}}$ no converge o converge a una matriz que no puede ser una generadora exacta de $\mathbf{P}$ puede aún tener una generadora exacta. Para la mayoría de las matrices de transición en la práctica esto será cierto, así se asume que las generadoras con suma de renglones de cero y que satisfaga $\mathbf{P} = \exp(\tilde{\mathbf{\Lambda}})$ pueden ser encontradas.

El principal inconveniente del teorema 6 es que la matriz $\tilde{\mathbf{\Lambda}}$ no garantiza tener elementos negativos fuera de la diagonal. Esto origina que $\mathbf{P}_t$ no sea una matriz de transición apropiada.

Sin embargo, algunas entradas negativas fuera de la traza de $\tilde{\mathbf{\Lambda}}$ serán usualmente muy pequeñas. Por consiguiente, es posible corregir el problema simplemente ajustando esta matriz usando uno de los siguientes métodos:

Ajuste Diagonal (AD)

Reemplazar las entradas negativas por ceros, y sumando el valor apropiado en la correspondiente entrada diagonal. Es decir,

$$\hat{\lambda}_{ij} = \begin{cases} 0 & \text{si } (i \neq j) \text{ y } \lambda_{ij} < 0 \\ \tilde{\lambda}_{ij} & \text{en cualquier otro caso} \end{cases}$$

para $i, j \in E$.

Además, el conjunto de los elementos que pertenecen a la traza son igual a la suma de los elementos no negtivos, es decir,

$$\hat{\lambda}_{ii} = -\sum_{j \neq i} \hat{\lambda}_{ij}$$

para $i \in E$.

Ajuste de Valores (AV)

Una versión diferente de $\hat{\mathbf{\Lambda}}$ puede ser obtenida sumando los valores negativos en todas las entradas del mismo renglón (no sólo en la entrada de la diagonal) que tiene el signo correcto, proporcional a sus valores absolutos, es decir, sea

$$G_i = |\tilde{\lambda}_{ii}| + \sum_{j \neq i} \text{máx}(\tilde{\lambda}_{ij}, 0); \qquad B_i = \sum_{i \neq j} \text{máx}(-\tilde{\lambda}_{ij}, 0)$$

así, se tiene que

$$\hat{\lambda}_{ij} = \begin{cases} 0 & \text{si } (i \neq j) \text{ y } \tilde{\lambda}_{ij} < 0 \\ \tilde{\lambda}_{ij} - |\tilde{\lambda}_{ij}| \frac{B_i}{G_i} & \text{en cualquier otro caso si } G_i > 0 \\ \tilde{\lambda}_{ij} & \text{en cualquier otro caso si } G_i = 0 \end{cases}$$

Cuando $\tilde{\mathbf{\Lambda}}$ tiene sólo pocos elementos negativos fuera de la traza, y los valores de $-\tilde{\lambda}_{ii}$ son razonablemente grandes los dos métodos anteriores serán muy parecidos.

Los teoremas 4 y 5 proporcionan un poco de validez a la matriz $\tilde{\mathbf{\Lambda}}$ del teorema 6. En particular de ellos se muestra inmediatemente lo siguiente

Corolario 1 *Sea $\boldsymbol{P}$ una matriz de transición tal que a lo más una de ls siguientes tres condiciones vale:*

- *$det(\boldsymbol{P}) > 1/2$ y $||\boldsymbol{P} - \boldsymbol{I}|| < 1/2$, o*
- *$\boldsymbol{P}$ tiene eigenvalores distintos y $det(\boldsymbol{P}) > e^{-\pi}$, o*
- *$\boldsymbol{P}$ distintos eigenvalores reales.*

sugóngase además que la serie (8) converge a una matriz $\tilde{\mathbf{\Lambda}}$ con elementos negativos fuera de la diagonal. Entonces no existe un generador valido para $\boldsymbol{P}$.

Método por programación dinámica

Calcular el riesgo de crédito de los valores requiere el calculo de probabilidades de transición sobre intervalos menores a un año. El supuesto de tiempo homogéneo en este caso conduce a el problema de encontrar la matriz de transición $\mathbf{Q}$ tal que

$$\mathbf{Q}^t = \mathbf{P}, \qquad t > 1$$

donde $\mathbf{P}$ es una matriz de transición anual y t es el número de periodos por año (es decir, $t = 12$ para una matriz de transición mensual).

Se define el conjunto de matrices de transición, $MT(k)$, que consiste de todas las matrices de dimensión $k \times k$ que satisfacen la condición de matriz estocástica.

Calculando $\mathbf{Q} = \mathbf{P}^{1/t} = \exp\left(\frac{1}{t}\ln(\mathbf{P})\right)$ podría resultar en una matriz que tiene entradas negativas y, de esta manera, $\mathbf{Q}$ podría no pertenecer a el conjunto $MT(k)$. Nótese que si existe una generadora $\tilde{\mathbf{\Lambda}}$, satisfaciendo $e^{\tilde{\mathbf{\Lambda}}} = \mathbf{P}$, entonces $\tilde{\mathbf{Q}} = \exp\left(\frac{1}{t}(\tilde{\mathbf{\Lambda}})\right)$ pertenece al conjunto $MT(k)$.

Sin embargo, en la mayoría de los casos prácticos, la matriz de transición anual $\mathbf{P}$ no tiene una matriz generadora. Por lo tanto, se introduce el concepto de regularización que permite resolver este problema.

El problema de regularización puede ser descrito de la siguiente forma: encontrar una matriz de transición $\mathbf{Q}$ que, elevada a la potencia t, sea la más cercana a la matriz de transición anual $\mathbf{P}$. En términos matemáticos, este problema puede ser formalmente establecido como:

Problema 1 (MAMT) *Mejor aproximación a la matriz de transición anual.*

Encontrar $\tilde{\boldsymbol{Q}} \in MT(k)$ tal que

$$||\tilde{\boldsymbol{Q}}^t - \boldsymbol{P}|| = \min_{\boldsymbol{Q} \in MT(k)} ||\boldsymbol{Q}^t - \boldsymbol{P}||$$

donde $||\cdot||$ es una norma apropiada en el espacio de las matrices de $k \times k$.

Optimización de la matriz (OM)

Encontrar $\hat{\mathbf{Q}} \in MT(k)$ tal que

$$||\hat{\mathbf{Q}} - \mathbf{P}^{1/t}|| = \min_{\mathbf{Q} \in MT(k)} ||\mathbf{Q} - \mathbf{P}^{1/t}||$$

De esta manera, el problema OM encuentra la matriz de transición que es lo más cercana posible a la matriz de transición anual fraccionada.

Para resolver la optimización de matrices, se usa el hecho de que el conjunto de matrices de transición $MT(k)$, puede ser representado como un producto Cartesiano de k-dimensiones. Es decir, cada fila de la matriz de transición satisface la condición de matriz estocástica y así pertenecen a el vector k-dimensional, $v(k)$, definido como

$$v(k) = \left\{ (x_1, \ldots, x_k) \in \mathbb{R}^k, \quad \sum_{i=1}^{k} x_i = 1, \quad x_i \geq 0 \right\}$$

Además, nótese que $v(k)$ esta contenido en el hiperplano $h(k)$

$$h(k) = \left\{ (x_1, \ldots, x_k) \in \mathbb{R}^k, \quad \sum_{i=1}^{k} x_i = 1 \right\}$$

Se utiliza la norma Euclidiana para medir la distancia entre cualquiera dos puntos x y y en $\mathbb{R}^k$:

$$d(x, y) = \sqrt{\sum_{i=1}^{k} (y_i - x_i)^2}, \quad x, y \in \mathbb{R}^k$$

Entonces el problema OM puede esencialmente ser resuleto renglón por renglón. Esto es, el problema OM puede ser reducido a k independientes casos del problema de minimización de distancias: para un punto dado $p \in \mathbb{R}^k$ (es decir, un renglón de la matriz $\mathbf{P}^{1/t}$), encontrar $x^* \in v(k)$ tal que

$$d(p, x^*) = \min_{x \in v(k)} d(y, x)$$

El algoritmo que puede solucionar lo anterior fué sugerido por Merkoulovith (2000), y es de la siguiente forma:

Paso 1 Encontrar la proyección $\mathbf{b}$ del punto p en el hiperplano $h(k)$, donde $b_i = p_i - \lambda$, con

$$\lambda = \frac{1}{k} \left(\sum_{i=i}^{k} p_i - 1 \right).$$

Paso 2 Si todas las coordenadas de **b** son no negativas entonces se para; **b** es la solución al problema.

Paso 3 Sea $\hat{p} = \pi(b)$, donde π es una permutación que ordena las coordenadas de **b** en una sucesión descendiente.

Paso 4 Se calcula $C_n = \sum_{i=i}^{n} \hat{p}_i - n\hat{p}_n$ para $n = 1 \ldots, k$. Las sumas C_n satisfacen $0 \leq C_1 \leq C_2 \leq \cdots \leq C_k$.

Paso 5 Encontrar $n^* = \text{máx}\{n : n \geq 1,\ C_n \leq 1\}$.

Paso 6 Construir el vector $\hat{x} \in v(k)$ como sigue. Para toda $j > n^*$, $\hat{x}_j = 0$, y para $j \leq n^*$,

$$\hat{x}_j = \hat{p}_j + \frac{1}{n^*}\left(1 - \sum_{i=1}^{n^*} \hat{p}_i\right).$$

Paso 7 Aplicar la permutación inversa π^{-1} a $\hat{x}$; $\pi^{-1}(\hat{x})$ es la solución al problema.

Comparando $MAMT$ y OM se espera que $\tilde{\mathbf{Q}}$ y $\hat{\mathbf{Q}}$ sean lo más cercanas posible, por esta razón es natural llamar a $\hat{\mathbf{Q}}$ una casi-solución al problema $MAMT$.

Optimización de la generadora (OG)

El segundo método utiliza a las matrices generadoras como objeto de regularización. Se define el conjunto de matrices generadoras, $G(k)$, compuesta de todas las matrices de dimensión $K \times K$ tal que satisfacen (7).

Encontrar $\hat{\mathbf{\Lambda}} \in G(k)$ tal que

$$||\hat{\mathbf{\Lambda}} - \ln(\mathbf{P})|| = \min_{\mathbf{\Lambda} \in \mathbf{G(k)}} ||\mathbf{\Lambda} - \ln(\mathbf{P})||$$

El problema OG es diferente del problema OM en un sentido geométrico. Mientras el espacio de matrices de transición, $MT(k)$, es un producto cartesiano, el espacio de sus generadoras, $G(k)$, es un producto cartesiano de conos k-dimensionales. Cada renglón de una generadora tiene la propiedad (7). Permutando los elementos de un renglón, siempre se pueden representar como un punto en un cono estándar, $C(k)$, definido por

$$C(k) = \left\{(x_1, \ldots, x_k) \in \mathbb{R}^k,\ \sum_{i=1}^{k} x_i = 0,\ x_1 \geq 0,\ x_i \leq 0, \forall\ i \leq 2\right\}$$

Nótese que $C(k)$ esta contenido en el hiperplano

$$H(k) = \left\{(x_1, \ldots, x_k) \in \mathbb{R}^k,\ \sum_{i=1}^{k} x_i = 0\right\}$$

En una forma similar al problema OM, el problema OG puede ser resuelto renglón por renglón basándose en proyectar un punto $p \in \mathbb{R}^k$ en el cono definido anteriormente. Así, el problema OG puede ser reducido a k casos independientes del siguiente problema de minimización de distancias: para un punto dado $p \in \mathbb{R}^k$ (es decir, un renglón de la matriz $\ln(\mathbf{P})$), encontrar $g^* \in C(k)$ tal que

$$d(p, g^*) = \min_{g \in C(k)} d(p, g)$$

La solución óptima a lo antrior puede ser obtenida como sigue:

Paso 1 Si $\mathbf{b}$ es la proyección de $p \in \hat{h}(k)$ tal que $b_i = p_i - \lambda$, donde

$$\lambda = \frac{1}{k}\left(\sum_{i=1}^{k} p_i\right)$$

Paso 2 Si $\hat{p} = \pi(b)$, donde π es una permutación que ordena las coordenadas de $\mathbf{b}$ en una sucesión ascendiente.

Paso 3 Encontrar ℓ^*, el más pequeño entero $2 \leq \ell \leq k-1$ que satisface

$$(k - \ell + 1)\hat{p}_{\ell+1} \geq \hat{p}_\ell + \sum_{i=0}^{k-(\ell+1)} \hat{p}_{k-i}$$

Paso 4 Si $S = \{i : 2 \leq i \leq \ell^*\}$. Construir el vector $\hat{g} \in C(k)$ como sigue

$$\hat{g}_i = \begin{cases} 0 & i \in S \\ \hat{p}_i - \frac{1}{k-\ell^*+1}\sum_{j \notin S} \hat{p}_j & \text{en cualquier otro caso} \end{cases}$$

Paso 5 Aplicar la permutación inversa π^{-1} a $\hat{g}$; así $\pi^{-1}(\hat{g})$ es la solución a el problema.

Los problema $MAMT$ y OG son relacionados bajo el supuesto que $\exp\left(\frac{1}{t}\hat{\mathbf{\Lambda}}\right)$ es cercana a $\tilde{\mathbf{Q}}$, y así la matriz $\hat{\mathbf{\Lambda}}$ puede también ser vista como una casi-solución al problema $MAMT$.

Cuando $\mathbf{P}^{1/t}$ no es una matriz de transición válida, los problemas $MAMT$ y OM encuentran soluciones en $MT(k)$ que son las más cercanas a $\mathbf{P}^{1/t}$. Análogamente, cuando $\ln(\mathbf{P})$ no es una generadora válida, el problema OG encuentra la más cercana posible matriz generadora $\hat{\mathbf{\Lambda}}$. La exponencial de la matriz generadora entonces produce una matriz de transición válida que es cercana a $\mathbf{P}^{1/t}$.

Función de verosimilitud

Este método demuestra que un modelo continuo de Markov puede asimismo ser usado para analizar observaciones discretas, donde la calificación sólo ha sido observada a tiempo discreto. Proporciona dos métodos para estimar las intensidades de transición, es decir, la generadora de un proceso de Markov.

Sea $\mathbf{X}_i = \left\{X_t^i\right\}_{t\geq 0}$, $i = 1, \ldots, N$ procesos de Markov independientes con el mismo espacio de estados finito $E = \{1, \ldots, m\}$ y la misma matriz de intensidad $\mathbf{\Lambda} = \{\lambda_{\mathbf{ij}}\}$. En particular, λ_{ij} es la probabilidad de transición del estado i a el estado j, y $\lambda_{ii} = -\sum_{j\neq i}\lambda_{ij}$. En el contexto de riesgo de crédito, $\mathbf{X}_i$ es es la calificación crediticia historica de la ith compañía.

Si las $\mathbf{X}_i$ han sido observadas continuamente en el intervalo de tiempo $[0, T]$, la función de verosimilitud esta dada por

$$L^{(c)}(\mathbf{\Lambda}) = \prod_{\mathbf{k=1}}^{\mathbf{N}} \prod_{\mathbf{i=1}}^{\mathbf{m}} \prod_{\mathbf{j\neq i}} \lambda_{\mathbf{ij}}^{\mathbf{N_{ij}^k(T)}} \mathbf{e}^{-\lambda_{\mathbf{ij}}\mathbf{R_i^k(T)}}$$

donde $N_{ij}(t)$ es el número de transiciones del estado i al estado j en el intervalo de tiempo $[0, t]$ del proceso $\mathbf{X}_k$,

$$R_i^k(t) = \int_0^t \mathbf{I}\left\{\mathbf{X}_s^k = i\right\} ds$$

es el tiempo gastado en el estado i antes del tiempo t por el proceso $\mathbf{X}_k$, y

$$N_{ij} = \sum_{k=1}^{N} N_{ij}^k(t), \qquad R_i(t) = \sum_{k=1}^{N} R_i^k(t);$$

Es fácil ver que el estimador de máxima verosimilitud de $\mathbf{\Lambda}$ basado en datos continuos de calificaciones crediticias historicas se da para toda $i \neq j$ por

$$\hat{\lambda}_{ij}^{(c)}(T) = \frac{N_{ij}(T)}{R_i(T)},$$

obviamente, $\hat{\lambda}_{ii}^{(c)}(T) = -\sum_{j\neq i}\hat{\lambda}_{ij}(T)$ para $i = 1, \ldots, K$.

Sea $0 \leq t_1^i < \cdots < t_{n_i}^i \leq T$ los tiempos en los cuales el proceso $\mathbf{X}_i$ ha sido observado. La observación historica $Y_i^k = X_{t_i^k}^k$ es una cadena de Markov discreta, para la cual la matriz de transición al tiempo i es $\mathbf{P}^{\Delta_i^k}(\mathbf{\Lambda})$, donde $\Delta_i^k = t_{i+1}^k - t_i^k$ y

$$\mathbf{P}^t(\mathbf{\Lambda}) = \exp(\mathbf{t\Lambda}), \qquad \mathbf{t > 0}$$

Nótese que el proceso a tiempo discreto es no homogéneo cuando los puntos t_i^k no son equidistantes. La función de verosimilitud para datos discretos es estonces dada por

$$L^{(d)}(\mathbf{\Lambda}) = \prod_{\mathbf{k=1}}^{\mathbf{N}} \prod_{\mathbf{i=1}}^{\mathbf{n_k-1}} \mathbf{P}^{\mathbf{\Delta_i^k}}(\mathbf{\Lambda})_{\mathbf{x_i^k x_{i+1}^k}},$$

donde $x_i^k, \ldots, x_{n_k}^k$ denota los valores observados de $\mathbf{X}_k$, y donde $\mathbf{P}_t(\mathbf{\Lambda})_{\mathbf{ij}}$ denota la $ijth$ entrada de la matriz de transición $\mathbf{P}^t(\mathbf{\Lambda})$.

Un caso especial es cuando $\Delta_i^k = \Delta$. En este caso

$$L^{(d)}(\mathbf{\Lambda}) = \prod_{\mathbf{i=1}}^{\mathbf{m}} \prod_{\mathbf{j=1}}^{\mathbf{m}} \mathbf{P}^{\mathbf{\Delta}}(\mathbf{\Lambda})_{\mathbf{ij}}^{\mathbf{K_{ij}}},$$

donde K_{ij} denota el número de transiciones de i a j observado en toda $\mathbf{Y}_\ell = \left\{Y_k^\ell\right\}_{k=1,\ldots,n_\ell}$, $\ell = 1\ldots, N$. El estimador de máxima verosimilitud de $\mathbf{P} = \exp(t\mathbf{\Lambda})$ esta aquí dado por

$$\hat{\mathbf{P}}_{ij} = \frac{K_{ij}}{K_{i.}}$$

donde $K_{i.} = \sum_{j=1}^m K_{ij}$. Puede pasar que $\hat{\mathbf{P}}$ no sea la matriz exponencial de una matriz de intensidad, pero incluso en ese caso el estimador de máxima verosimilitud de $\mathbf{\Lambda}$ podría existir.

Métricas para matrices

Una tarea importante por evaluar la bondad de los modelos o comparar las predicciones con matrices de transición reales es medir la distancia entre dos matrices. Se pueden encontrar varias medidas basadas en distancia celda por celda, eigenvalores, eigenvectores, o métricas basadas en valores singulares. A continuación se expondrá un breve sumario de las medidas descritas hasta ahora.

Medidas de distancia clásica y de celda por celda

Para comparar las bondades de aproximaciones, es necesario calcular las distancias de la matriz de transición original a las de predicción o aproximación.

Probablemente las medidas de distancia más intuitivas son la de celda-por-celda las métricas $\mathcal{L}^1$ y $\mathcal{L}^2$. Las cuales pueden denotarse como

$$D_{\mathcal{L}^1}(\mathbf{P}, \mathbf{Q}) = \sum_i \sum_j |p_{ij} - q_{ij}|$$

$$D_{\mathcal{L}^2}(\mathbf{P}, \mathbf{Q}) = \sum_i \sum_j (p_{ij} - q_{ij})^2$$

donde $i, j \in E$.

La diferencias del valor absoluto y el cuadrado podrían ser agrupadas en la categoría norma de Hölder a la potencia p con p variando desde 1 a infinito:

$$D_{\mathcal{L}^p}(\mathbf{P}, \mathbf{Q}) = \sum_i \sum_j |p_{ij} - q_{ij}|^p$$

Otra medida de distancia sería usar la distancia máxima de todas las celdas

$$D_{\mathcal{L}^{\text{máx}}}(\mathbf{P}, \mathbf{Q}) = \max_{i,j} |p_{ij} - q_{ij}|$$

Aplicación

En finanzas, el espacio de estados finito $\Omega = \{1, 2, \ldots, K\}$ cubre a las posibles calificaciones de los créditos, con el estado 1 siendo la calificación más alta y el estado K siendo el incumplimiento. Típicamente,

$$\Omega = \{Aaa, Aa, A, Baa, Ba, B, C, D\}.$$

En la medida en que una calificación es muy baja existe mayor riesgo de crédito, una matriz de transición de calificaciones debe satisfacer una de las dos condiciones del Lema 1.

Se ilustrará lo anterior con un ejemplo. Considérese la siguiente matriz de transición anual.

P	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	0.9340	0.0594	0.0064	0	0.0002	0	0	0
Aa	0.0161	0.9055	0.0746	0.0026	0.0009	0.0001	0	0.0002
A	0.0007	0.0228	0.9244	0.0463	0.0045	0.0012	0.0001	0
Baa	0.0005	0.0026	0.0551	0.8848	0.0476	0.0071	0.0008	0.0015
Ba	0.0002	0.0005	0.0042	0.0516	0.8691	0.0591	0.0024	0.0129
B	0	0.0004	0.0013	0.0054	0.0635	0.8422	0.0191	0.0681
C	0	0	0	0.0062	0.0205	0.0408	0.6920	0.2405
D	0	0	0	0	0	0	0	1

Por el Teorema 3 (c), la matriz de transición $\mathbf{P}$ *no tiene una generadora exacta*, ya que $p_{23} > 0$ y $p_{37} > 0$, pero $p_{27} = 0$. Así, el Teorema 4 (a), indica que la matriz $\mathbf{P}$ tiene a lo más una matriz generadora (puesto que $det(\mathbf{P}) = 0{,}3448 > 0{,}5$), por (b) se sabe que esta generadora no puede ser $Ln(\mathbf{P})$ ya que $||\mathbf{P} - \mathbf{I}|| = 0{,}5118 > 0{,}5$.

ln(**P**)	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	-0.0688	0.0646	0.0043	-0.0002	0.0002	-0.0000	-0.0000	-0.0000
Aa	0.0175	-0.1009	0.0816	0.0008	0.0008	0.0000	-0.0000	0.0002
A	0.0005	0.0249	-0.0812	0.0511	0.0036	0.0010	0.0001	-0.0001
Baa	0.0005	0.0021	0.0608	-0.1256	0.0540	0.0063	0.0009	0.0009
Ba	0.0002	0.0004	0.0029	0.0587	-0.1445	0.0690	0.0022	0.0111
B	-0.0000	0.0004	0.0012	0.0040	0.0740	-0.1750	0.0249	0.0705
C	-0.0000	-0.0000	-0.0003	0.0070	0.0242	0.0525	-0.3689	0.2855
D	0	0	0	0	0	0	0	0

En consecuencia, será útil encontrar una matriz generadora aproximada $\mathbf{\Lambda}$, tal que $\exp(\mathbf{\Lambda})$ sea aproximadamente igual a $\mathbf{P}$, y que una matriz de transición para algún tiempo t pueda ser aproximada por $\exp(t\mathbf{\Lambda})$.

En base a los resultados anteriores, se procede a estimar una matriz generadora $\mathbf{\Lambda}$ con los distintos modelos expuestos anteriormente.

Modelo JLT

El modelo JLT proporciona una generadora aproximada $\mathbf{\Lambda_{JLT}}$, asumiendo que cada empresa hizo a lo más una transacción a lo largo del periodo. Así su generadora aproximada es

$\mathbf{\Lambda_{JLT}}$	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	-0.0683	0.0615	0.0066	0	0.0002	0	0	0
Aa	0.0169	-0.0993	0.0784	0.0027	0.0009	0.0001	0	0.0002
A	0.0007	0.0237	-0.0786	0.0481	0.0047	0.0012	0.0001	0
Baa	0.0005	0.0028	0.0585	-0.1224	0.0506	0.0075	0.0008	0.0016
Ba	0.0002	0.0005	0.0045	0.0553	-0.1403	0.0633	0.0026	0.0138
B	0	0.0004	0.0014	0.0059	0.0691	-0.1717	0.0208	0.0741
C	0	0	0	0.0074	0.0245	0.0488	-0.3682	0.2875
D	0	0	0	0	0	0	0	0

Nótese que en $\mathbf{\Lambda_{JLT}}$ los renglones suman cero y no hay entradas negativas fuera de la diagonal, lo cual indica que es una matriz generadora valida. Por lo anterior, si se calcula $\exp(\mathbf{\Lambda_{JLT}})$ se tiene que

$\exp(\mathbf{\Lambda_{JLT}})$	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	0.9345	0.0566	0.0084	0.0003	0.0002	0	0	0
Aa	0.0156	0.9068	0.0719	0.0042	0.0011	0.0002	0	0.0002
A	0.0009	0.0218	0.9266	0.0437	0.0053	0.0014	0.0001	0.0001
Baa	0.0005	0.0031	0.0532	0.8873	0.0448	0.0080	0.0008	0.0022
Ba	0.0002	0.0006	0.0056	0.0488	0.8723	0.0545	0.0025	0.0154
B	0	0.0004	0.0016	0.0068	0.0595	0.8445	0.0160	0.0711
C	0	0	0.0003	0.0065	0.0206	0.0380	0.6924	0.2422
D	0	0	0	0	0	0	0	1

Así, se observa que $\exp(\mathbf{\Lambda_{JLT}})$ es cercana a, pero no exactamente igual a, la origanl $\mathbf{P}$.

Modelo IRW

En el modelo IRW se usa el método presentado en el Teorema 6. De hecho, se puede observa que para cuatro cifras decimales la serie (8) para $\mathbf{\Lambda}$ converge muy rápidamente, obteniendose una matriz generadora aproximada,

Si se calcula la matriz exponencial de la matriz anterior, se podría observar que $\exp(\mathbf{\Lambda_{IRW}}) = \mathbf{P}$ exactamente. Además, los renglones de $\mathbf{\Lambda_{IRW}}$ suman cero,

$\mathbf{\Lambda_{IRW}}$	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	-0.0688	0.0646	0.0043	-0.0002	0.0002	-0.0000	-0.0000	-0.0000
Aa	0.0175	-0.1009	0.0816	0.0008	0.0008	0.0000	-0.0000	0.0002
A	0.0005	0.0249	-0.0812	0.0511	0.0036	0.0010	0.0001	-0.0001
Baa	0.0005	0.0021	0.0608	-0.1256	0.0540	0.0063	0.0009	0.0009
Ba	0.0002	0.0004	0.0029	0.0587	-0.1445	0.0690	0.0022	0.0111
B	-0.0000	0.0004	0.0012	0.0040	0.0740	-0.1750	0.0249	0.0705
C	-0.0000	-0.0000	-0.0003	0.0070	0.0242	0.0525	-0.3689	0.2855
D	0	0	0	0	0	0	0	0

como debe ser. Desafortunadamente, $\mathbf{\Lambda_{IRW}}$ tiene algunas entradas negativas fuera de la diagonal. Si se reemplaza $\mathbf{\Lambda_{IRW}}$ por $\tilde{\mathbf{\Lambda}}_{IRW}$ en la cual se hace un ajuste (AD), se obtiene una matriz generadora valida

$\tilde{\mathbf{\Lambda}}_{IRW}$	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	-0.0690	0.0646	0.0043	0	0.0002	0	0	0
Aa	0.0175	-0.1009	0.0816	0.0008	0.0008	0.0000	0	0.0002
A	0.0005	0.0249	-0.0813	0.0511	0.0036	0.0010	0.0001	0
Baa	0.0005	0.0021	0.0608	-0.1256	0.0540	0.0063	0.0009	0.0009
Ba	0.0002	0.0004	0.0029	0.0587	-0.1445	0.0690	0.0022	0.0111
B	0	0.0004	0.0012	0.0040	0.0740	-0.1750	0.0249	0.0705
C	0	0	0	0.0070	0.0242	0.0525	-0.3692	0.2855
D	0	0	0	0	0	0	0	0

Así, $\tilde{\mathbf{\Lambda}}_{IRW}$ no tiene entradas negativas fuera de la diagonal y aún los renglones suman cero. Sin embargo, no satisface exactamente $\exp(\tilde{\mathbf{\Lambda}}_{IRW}) = \mathbf{P}$. De hecho, se muestra la matriz de transición $\exp(\tilde{\mathbf{\Lambda}}_{IRW})$ la cual es muy cercana pero no exactamente igual a, la matriz de transición $\mathbf{P}$ original.

$\exp(\tilde{\mathbf{\Lambda}}_{IRW})$	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	0.9338	0.0594	0.0064	0.0002	0.0002	0	0	0
Aa	0.0161	0.9055	0.0746	0.0026	0.0009	0.0001	0	0.0002
A	0.0007	0.0228	0.9243	0.0463	0.0045	0.0012	0.0001	0.0001
Baa	0.0005	0.0026	0.0551	0.8848	0.0476	0.0071	0.0008	0.0015
Ba	0.0002	0.0005	0.0042	0.0516	0.8691	0.0591	0.0024	0.0129
B	0.0000	0.0004	0.0013	0.0054	0.0635	0.8422	0.0191	0.0681
C	0.0000	0.0000	0.0002	0.0062	0.0205	0.0408	0.6918	0.2405
D	0	0	0	0	0	0	0	1

Prgramación Dinámica

En el método de Programación Dinámica se utiliza la optimización de matrices generadoras (OG), es decir, se calcula $Ln(\mathbf{P})$ como en la mayoría de las ocasiones no es una matriz generadora valida, se aplica el problema de regularización, el cual puede ser resuelto por (OG), así con este método se obtiene la siguiente matriz

$\mathbf{\Lambda_{OG}}$	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	-0.0689	0.0645	0.0042	0	0.0001	0	0	0
Aa	0.0175	-0.1009	0.0816	0.0008	0.0008	0.0000	0	0.0002
A	0.0005	0.0249	-0.0812	0.0511	0.0036	0.0010	0.0001	0
Baa	0.0005	0.0021	0.0608	-0.1256	0.0540	0.0063	0.0009	0.0009
Ba	0.0002	0.0004	0.0029	0.0587	-0.1445	0.0690	0.0022	0.0111
B	0	0.0004	0.0012	0.0040	0.0740	-0.1750	0.0249	0.0705
C	0	0	0	0.0070	0.0241	0.0524	-0.3690	0.2854
D	0	0	0	0	0	0	0	0

Se observa que $\mathbf{\Lambda_{OG}}$ no tiene entradas negativas fuera de la diagonal y los renglones suman cero, por lo tanto es considerada una matriz generadora valida. Sin embargo, no satisface exactamente $\exp(\mathbf{\Lambda_{OG}}) = \mathbf{P}$, ya que al calcular $\exp(\mathbf{\Lambda_{OG}})$ se obtiene la siguiente matriz.

$\exp(\mathbf{\Lambda_{OG}})$	Aaa	Aa	A	Baa	Ba	B	C	D
Aaa	0.9340	0.0594	0.0064	0.0002	0.0002	0.0000	0.0000	0.0000
Aa	0.0161	0.9055	0.0746	0.0026	0.0009	0.0001	0.0000	0.0002
A	0.0007	0.0228	0.9244	0.0463	0.0045	0.0012	0.0001	0.0001
Baa	0.0005	0.0026	0.0551	0.8848	0.0476	0.0071	0.0008	0.0015
Ba	0.0002	0.0005	0.0042	0.0516	0.8691	0.0591	0.0024	0.0129
B	0.0000	0.0004	0.0013	0.0054	0.0635	0.8422	0.0191	0.0681
C	0.0000	0.0000	0.0002	0.0062	0.0204	0.0407	0.6920	0.2404
D	0	0	0	0	0	0	0	1

la cual es muy cercana pero no exactamente igual a, la matriz de transición $\mathbf{P}$ original.

Una vez obtenidas las estimaciones de $\mathbf{\Lambda}$ se utilizarán distintas metricas para saber cual es la mejor estimación, y así poder valorar que método es el adecuado para el calculo de la matriz generadora valida, esto se mustra en la siguiente tabla.

	$\mathbf{\Lambda_{JLT}}$	$\tilde{\mathbf{\Lambda}}_{IRW}$	$\mathbf{\Lambda_{OG}}$
$D_{\mathcal{L}^1}(\mathbf{\Lambda}, \cdot)$	0.0771	0.0012	0.0014
$D_{\mathcal{L}^2}(\mathbf{\Lambda}, \cdot)$	0.00023855	0.00000028	0.00000022
$D_{\mathcal{L}^{\text{máx}}}(\mathbf{\Lambda}, \cdot)$	0.0057	0.0003	0.0003

Cuadro 1: Comparación de matrices generadoras

Análogamente, para las matrices de transición estimadas con los distintos modelos.

	$\mathbf{P}_{JLT}$	$\tilde{\mathbf{P}}_{IRW}$	$\mathbf{P}_{OG}$
$D_{\mathcal{L}^1}(\mathbf{P},\cdot)$	0.0628	0.0010	0.0008
$D_{\mathcal{L}^2}(\mathbf{P},\cdot)$	0.0001571	0.00000018	0.00000012
$D_{\mathcal{L}^{\text{máx}}}(\mathbf{P},\cdot)$	0.0046	0.0002	0.0002

Cuadro 2: Comparación de matrices de transición

En base a los resultados obtenidos en las Tablas 1 y 2, se puede concluir que el mejor de los métodos vistos para estimar matrices generadoras es Programación Dinámica, ya que la bondad de este modelo es la menor, tanto para la matriz generadora como la matriz de transición estimadas.

Anexo A

Fundamentos de probabilidad

Un espacio medible $(\Omega, \mathcal{F})$ consiste de un espacio muestra Ω y una colección de subconjuntos de Ω llamada una σ-álgebra. Un mapeo contable aditivo $\mu : \mathcal{F} \rightarrow \mathbb{R}_+$ es llamado una medida en $(\Omega, \mathcal{F})$ y la tripleta $(\Omega, \mathcal{F}, \mu)$ es llamada un espacio medible.

Una función $\mathcal{F}$-medible es un mapeo h de $(\Omega, \mathcal{F})$ a un espacio de estados $(\mathcal{S}, \mathscr{S})$ tal que para algún $\mathcal{A} \in \mathscr{S}$, $h^{-1}(\mathcal{A}) \in \mathcal{F}$. Una σ-álgebra de Borel es una *sigma*-álgebra que contiene a todos los conjuntos abiertos de un espacio topológico.

Un *proceso* es una familia de variables aleatorias y se escribe $(X_t)_{0 \leq t \leq T}$ o simplemente X_t. El índice t es frecuentemente interpretado como el tiempo, definido en el intervalo $[0, T]$ para $T < \infty$.

Para los puntos de la muestra fijos $\omega \in \Omega$, la función $t \rightarrow X_t(\omega)$ es llamada la *trayectoria* del proceso. No obstante, un proceso puede ser visto como una fucnión de dos variables (t, ω) donde t es el (tiempo) índice y $\omega \in \Omega$, es decir, se tiene el mapeo $(t, \omega) \rightarrow X_t(\omega)\ \forall \omega \in \Omega$. Por consiguiente, un proceso es un mapeo conjunto h de $\mathbb{R}_+ \times \Omega$ en $mathcalS$. Este mapeo es *medible* si se proporciona la correspondiente *sigma*-álgebra de los conjuntos de Borel, es decir, si se tiene $h : (t, \omega) \rightarrow X_t(\omega)\ :\ (\mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F} \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)))$ tal que para algún $\mathcal{A} \in \mathcal{B}(\mathbb{R}^d)$, $h^{-1}(\mathcal{A}) \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$.

Una *filtración* $\mathcal{F}_{0 \leq t \leq T} = \{\mathcal{F}_0, \ldots, \mathcal{F}_T\}$, algunas veces escrito como $(\mathcal{F}_t)_{0 \leq t \leq T}$ o $\mathcal{F}_t$, es una familia de sub-σ-álgebra incluida en $\mathcal{F}$ la cual es no decreciente en el sentido de que $\mathcal{F}_s \subseteq \mathcal{F}_t$ para $s < t$. Incluso, una filtración puede ser interpretada como representando la acumulación de información sobre el tiempo. Un proceso genera una filtración.

Un espacio de probabilidad filtrado es un espacio de probabilidad equipado con una filtración y es a veces escrito como $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbf{P})$, $(\Omega, \mathcal{F}, (\mathcal{F}_{t \in [0,T]}), \mathbf{P})$, o $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$.

Sea un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbf{P})$ dado. Un proceso $(X_t)_{0 \leq t \leq T}$ es llamado *adaptado* (a la filtración $\mathcal{F}_t$) si es $\mathcal{F}_t$-medible. TAmbién se dice que el proceso es $\mathcal{F}_t$-adaptado.

Clases de procesos

Definición 1 *Sea $\mathcal{L}^\infty$ la clase de todos los procesos adaptados, si θ es un proceso. Se definen los siguientes espacios:*

$$\mathcal{L}^1 = \left\{ \theta \in \mathcal{L}^\infty : \int_0^T |\theta(t)| dt < \infty \quad c.s. \right\}$$

$$\mathcal{L}^2 = \left\{ \theta \in \mathcal{L}^\infty : \int_0^T \theta^2(t)dt < \infty \quad c.s. \right\}$$

$$\mathcal{H}^2 = \left\{ \theta \in \mathcal{L}^2 : \mathbb{E}\left(\int_0^T \theta^2(t)dt \right) < \infty \right\}$$

$$\mathcal{L}^{1,1} = \left\{ \theta \in \mathcal{L}^\infty : \int_0^{T_1} \int_0^{T_2} |\theta(s,t)| dsdt < \infty \quad c.s. \right\}$$

$$\mathcal{L}^{2,1} = \left\{ \theta \in \mathcal{L}^{1,1} : \int_0^T \theta^2(s,u)ds < \infty \quad c.s., \quad \int_0^{T_1} \left(\int_0^{T_2} |\theta(s,u)| du \right)^2 ds < \infty \quad c.s. \right\}$$

Se escibirá $\theta \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbf{P})$ o simplemente $\theta \in \mathcal{L}^1$ si es claro del contexto en que espacio de probabilidad el proceso es definido.

Martingalas

Definición 2 *Un proceso estocástico $M(t)$[1] adaptado a una filtración $\mathbb{F} = (\mathcal{F}_t)$ es una martingala si para alguna t, $M(t)$ es integrable, esto es, $\mathbb{E}(|M(t)|) < \infty$ y para alguna t y s con $0 \leq s \leq t \leq T$,*

$$\mathbb{E}\left(M(t)|\mathcal{F}_s\right) = M(s) \quad \textit{casi seguramente} \tag{1}$$

$M(t)$ es una martingala en $[0, \infty)$ si es integrable y (1) vale para alguna $0 \leq s \leq t \leq \infty$.

Definición 3 *Una variable aleatoria X es cuadrado integrable si $\mathbb{E}(X^2) < \infty$. Un proceso $X(t),\ t \geq 0$ es cuadrado integrable si $\sup_{t \geq 0} \mathbb{E}(X^2(t)) < \infty$. Si $X(t)$ es considerada en un intervalo finito $0 \leq t \leq T$, entonces es cuadrado integrable si $\sup_{0 \leq t \leq T} \mathbb{E}(X^2(t)) < \infty$. Si $X(t)$ es una martingala y es cuadrado integrable, entonces es llamada martingala cuadrado integrable.*

Un proceso $X(t)$ de variación cuadrática está definido como un límite en probabilidad

$$[X, X](t) = \lim \sum_{i=1}^{n} \left(X(t_i^n) - X(t_{i-1}^n) \right)^2 \tag{2}$$

donde el límite se toma sobre la partición

$$0 = t_0^n < t_1^n < \ldots < t_n^n = t.$$

Teorema 1 *Sea $M(t)$ una martingala con los segundos momentos finitos, es decir, $\mathbb{E}(M^2(t)) < \infty$ para toda t. Entonces sus proceso de variación cuadrática $[M, M](t)$ definido en (2) existe, y además $M^2(t) - [M, M](t)$ es una martingala.*

[1] Donde el tiempo t es continuo $0 \leq t \leq T$ o discreto $t = 0, 1, \ldots, T$

Medidas de probabilidad equivalentes

Sean dos medidas de probabilidad $\mathbf{P}$ y $\mathbf{Q}$ definidas en el mismo espacio. $\mathbf{P}$ y $\mathbf{Q}$ son *equivalentes* si ellas tienen los mismos conjuntos nulos.

Definición 4 *$\boldsymbol{Q}$ es absolutamente continua con respecto a $\boldsymbol{P}$, si $\boldsymbol{Q}(A) = 0$ siempre que $\boldsymbol{P}(A) = 0$ y se denota como $\boldsymbol{Q} \ll \boldsymbol{P}$. Por otro lado $\boldsymbol{P}$ y $\boldsymbol{Q}$ son equivalentes si $\boldsymbol{Q} \ll \boldsymbol{P}$ y $\boldsymbol{P} \ll \boldsymbol{Q}$.*

Teorema 2 (Teorema de Girsanov) *Sea $\boldsymbol{Q} \ll \boldsymbol{P}$, entonces existe una variable aleatoria Λ, tal que $\Lambda \geq 0$, $\mathbb{E}_{\boldsymbol{P}}(\Lambda) = 1$, y*

$$\boldsymbol{Q}(A) = \mathbb{E}_{\boldsymbol{P}}(\Lambda \boldsymbol{1}(A)) = \int_A \Lambda d\boldsymbol{P} \tag{3}$$

para algún conjunto medible A. Λ es $\boldsymbol{P}$ casi ciertamente única. Recíprocamente, si existe Λ con las propiedades anteriores y $\boldsymbol{Q}$ esta definido (3), entonces es una medida de probabilidad y $\boldsymbol{Q} \ll \boldsymbol{P}$.

La variable aleatoria Λ en el teorema anterior es conocida como la *derivada de Radon-Nikodym* o la densidad de $\mathbf{Q}$ con respecto a $\mathbf{P}$, y se denota $d\mathbf{Q}/d\mathbf{P}$. Se sigue de (3) que si $\mathbf{Q} \ll \mathbf{P}$, entonces las esperanzas bajo $\mathbf{P}$ y $\mathbf{Q}$ están relacionadas por

$$\mathbb{E}_{\mathbf{Q}}(Z) = \mathbb{E}_{\mathbf{P}}(\Lambda Z) \tag{4}$$

para alguna variable aleatoria Z integrable con respecto a $\mathbf{Q}$.

Corolario 1 *Dado un espacio de probabilidad filtrado, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \boldsymbol{P})$, se define a $\gamma(t) = (\gamma_t^1, \ldots, \gamma_T^d)$ como un proceso adaptado en $(\mathcal{L}^2)^d$ tal que*

$$\frac{d\boldsymbol{Q}}{d\boldsymbol{P}} = \exp\left(\int_0^T \gamma(s)dW(s) - \frac{1}{2}\int_0^T ||\gamma(s)||^2 ds\right) \tag{5}$$

y $W(t) = (W_t^1, \ldots, W_t^d)$ un movimiento Browniano en $\mathbb{R}^d$. Sí se asume que la condición de Novikov es valida tal que $\zeta(t) = \mathbb{E}[\frac{d\boldsymbol{Q}}{d\boldsymbol{P}}]$ es una martingala. Entonces el proceso

$$\tilde{W}(t) = W(t) - \int_0^t \gamma(s)ds$$

es un movimiento Browniano d-dimensional definido en $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \boldsymbol{Q})$.

Cambio de medida para procesos

Hay un cambio de medida para el movimiento Browniano con drift que lo hace en un movimiento Browniano sin drift, similar a restar la media a variables aleatorias Gausianas.

Puesto que las construcciones de un movimiento Browniano son funciones continuas, las medidas de probabilidad son definidas en conjuntos medibles de funciones continuas en $[0, T]$, es decir, $C([0, T])$. Si $\omega(t)$ denota una función continua, entonces existe una medida de probabilidad $\mathbf{P}$ en este espacio tal que el proceso de la coordenada $B(t, \omega) = \omega(t)$ es un movimiento Browniano. $\mathbf{P}$ [2] es la medida de Wiener. Si $\mathbf{P}$ y $\mathbf{Q}$ son equivalentes, entonces existe una variable aleatoria $\Lambda = d\mathbf{Q}/d\mathbf{P}$, tal que las probabilidades bajo $\mathbf{Q}$ están dadas por $\mathbf{Q}(A) = \int_A \Lambda d\mathbf{P}$.

Teorema 3 *Sea $\Lambda(t)$ una martingala positiva con respecto a $\boldsymbol{P}$ tal que $\mathbb{E}_{\boldsymbol{P}}(\Lambda(T)) = 1$. Se define una nueva medida de probabilidad $\boldsymbol{Q}$ dada por $\boldsymbol{Q}(A) = \int_A \Lambda d\boldsymbol{P}$. Entonces $\boldsymbol{Q}$ es absolutamente continua con respecto a $\boldsymbol{P}$ y para alguna variable aleatoria X*

$$\mathbb{E}_{\boldsymbol{Q}}(X) = \mathbb{E}_{\boldsymbol{P}}(\Lambda(T)X)$$

$$\mathbb{E}_{\boldsymbol{Q}}(X|\mathcal{F}_t) = \mathbb{E}_{\boldsymbol{P}}(\frac{\Lambda(T)}{\Lambda(t)}X|\mathcal{F}_t)$$

y si X es medible con respecto a $\mathcal{F}_t$, entonces para $s \leq t$

$$\mathbb{E}_{\boldsymbol{Q}}(X|\mathcal{F}_s) = \mathbb{E}_{\boldsymbol{P}}(\frac{\Lambda(t)}{\Lambda(s)}X|\mathcal{F}_s) \tag{6}$$

Se sigue inmediatamente de (6) que

Corolario 2 *Un proceso $W(t)$ es una martingala con respecto a $\boldsymbol{Q}$ si y sólo si $\Lambda(t)W(t)$ es una martingala con respecto a $\boldsymbol{P}$.*

Teorema 4 *Sea $B(t)$, $0 \leq t \leq T$ un movimiento Browniano bajo la medida de probabilidad $\boldsymbol{P}$, y $\mu \neq 0$. Considérese el proceso $W(t) = B(t) + \mu t$. Existe una medida $\boldsymbol{Q}$ equivalente a $\boldsymbol{P}$, tal que $W(t)$ es un movimiento Browniano con respecto a $\boldsymbol{Q}$. Λ esta dada por*

$$\Lambda = \frac{d\boldsymbol{Q}}{d\boldsymbol{P}}(B) = e^{-\mu B(T) - \frac{1}{2}\mu^2 T} \quad y \quad \frac{1}{\Lambda} = \frac{d\boldsymbol{P}}{d\boldsymbol{Q}}(W) = e^{\mu W(T) - \frac{1}{2}\mu^2 T} \tag{7}$$

---

[2] Sea $\Omega = C([0, T], \mathbb{R})$ un espacio métrico medible y completo con respecto a la norma $d_1(f, g) = \sup_{t \in T} ||f(t) - g(t)||$, entonces a la medida $\mathbf{P}$ sobre $\mathcal{B}(\Omega)$ se llama la medida de Wiener, donde el espacio de probabilidad $(\Omega, \mathcal{B}(\Omega), \mathbf{P})$ se llama el espacio canónico del Mov. Browniano.

Demostración

Para la demostración se usara la caracterización de Lévy de un movimiento Browniano, como una martingala continua con proceso de variación cuadrática t. La variación cuadrática es una propiedad de trayectoria, no depende de la medida de probabilidad, con tal de que sea absolutamente continua con respecto a una dada. Por consiguente y usando el hecho de que μt es uniforme y no contribuye a la variación cuadratica.

$$[W,W](t) = [B(t)+\mu t, B(t)+\mu t] = [B,B](t) = t$$

Esto permanece estable cuando $W(t)$ es martingala bajo $\mathbf{Q}$. Por otro lado sea $\Lambda(t) = \mathbb{E}_{\mathbf{P}}(\Lambda|\mathcal{F}_t)$. Por el Corolario (2) y el Teorema (3) se muestra que $\Lambda(t)W(t)$ es una martingala bajo $\mathbf{P}$,

$$\mathbb{E}_{\mathbf{P}}\left(W(t)\Lambda(t)|\mathcal{F}_s\right) = \mathbb{E}_{\mathbf{P}}\left((B(t)+\mu t)e^{-\mu B(t)-\frac{1}{2}\mu^2 t}\right) = W(s)\Lambda(s)$$

Esto se hace con cálculos directos. □

Resulta que no sólo la tendencia lineal puede quitarse por un cambio de medida, sino cualquier tendencia de la forma $\int_0^t H(s)dB(s)$ con un proceso predecible H.

Cambio de medida de Wiener

Se demostrará que la derivada de Radon-Nikodym de cualquier cambio equivalente de medida de Wiener es una exponencial estocástica de algún proceso predecible. Usándose la propiedad de la representación predecible de la martingala Browniana se demuestran los siguientes resultados.

Teorema 5 *Sea $M(t)$, $0 \le t \le T$, una martingala local adaptada a un Browniano filtrado $\boldsymbol{F}$. Entonces existe un proceso predecible $H(t)$ tal que*

$$\mathbb{P}\left(\int_0^T H^2(s)ds < \infty\right) = 1$$

y

$$M(t) = M(0) + \int_0^t H(s)dB(s) \tag{8}$$

Además, si Y es una variable aleatoria integrable, $\mathbb{E}(Y) < \infty$, entonces $Y = \mathbb{E}\left[Y \exp\left(\int_0^T q(t)dB(t) - \frac{1}{2}\int_0^T q^2(t)dt\right)\right]$

Teorema 6 *Sea $\boldsymbol{F}$ un movimiento Browniano filtrado y sea Y una variable aleatoria positiva con esperanza finita, $\mathcal{F}_T$-medible. Entonces existe un proceso predecible $q(t)$ tal que*

$$Y = \mathbb{E}\left[Y \exp\left(\int_0^T q(t)dB(t) - \frac{1}{2}\int_0^T q^2(t)dt\right)\right]$$

Demostración
Sea $M(t) = \mathbb{E}(Y|\mathcal{F}_t)$. Entonces $M(t), 0 \leq t \leq T$ es una martingala integrable uniformemente positiva. Por el teorema 5 $M(t) = \mathbb{E}\left(Y + \int_0^t H(s)dB(s)\right)$. Se define a $q(t) = H(t)/M(t)$. Obsérvese que $q(t)$ esta propiamente definido, se mostrará que $M(t)$ nunca se hace cero en $[0,T]$. Sea

$$\tau = \text{ínf}\,\{t \in [0,T] : M(t) = 0\} \wedge T.$$

Entonces τ es un tiempo de paro acotado. Por el paro optativo se tiene que $\mathbb{E}[M(\tau)] = \mathbb{E}[M(T)]$. Pero

$$\mathbb{E}[M(\tau)] = \mathbb{E}[M(T)\mathbf{1}(\tau = T)] + \mathbb{E}[M(T)\mathbf{1}(\tau < T)] = \mathbb{E}[M(T)\mathbf{1}(\tau = T)].$$

Por consiguiente, $\mathbb{E}[M(T)(1 - \mathbf{1}(\tau = T))] = 0$. Puesto que $M(T)(1 - \mathbf{1}(\tau = T)) \geq 0$ casi seguramente, de eso se sigue que $\mathbb{P}(\tau = T) = 1$. Por lo tanto se puede escribir

$$dM(t) = H(t)dB(t) = M(t)q(t)dB(t) = M(t)dX(t) \tag{9}$$

y $M(t)$ es una semimartingala exponencial de $X(t) = \int_0^t q(s)dB(s)$, y se sigue el resultado. □

Corolario 3 *Sea $\boldsymbol{P}$ la medida de Wiener, $B(t)$ es un movimiento Browniano bajo $\boldsymbol{P}$, y $\boldsymbol{Q}$ es equivalente a $\boldsymbol{P}$. Entonces existe un proceso predecible $q(t)$, tal que*

$$\Lambda = \frac{d\boldsymbol{Q}}{d\boldsymbol{P}} = \exp\left(\int_0^T q(t)dB(t) - \frac{1}{2}\int_0^T q^2(t)dt\right) \tag{10}$$

Demostración
Ya que $\mathbb{P}(\Lambda > 0) = 1$ y $\mathbb{E}(\Lambda) = 1$, entonces se sigue el resultado por el Teorema (6) □

Integración estocástica

Teorema 7 (Doléans-Dade) *Sea $M(t)$ una martingala local continua definida en $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \boldsymbol{P})$. Si un proceso $\Lambda_t(M)$ satisface la ecuación diferencial estocástica*

$$d\Lambda(t) = \Lambda(t)dM(t),$$

entonces $\Lambda_t(M)$ está dada por la martingala local

$$\Lambda_t(M) = \exp\left(M(t) - \frac{1}{2}\langle M\rangle_t\right)$$

donde $\Lambda(t)$ es llamada la exponencial Doléans-Dade o exponencial estocástica.

Corolario 4 (Novikov) *Si $\mathbb{E}\left[\exp\left(\frac{1}{2}\langle M\rangle\right)\right] < \infty$ entonces Λ en el teorema (7) es una martingala.*

Anexo B

Progamación en Matlab

Aquí se presenta el algoritmo de los programas realizados como ayuda para el desarrollo de este trabajo.

Los siguientes progrmas verifican si se cumplen las condiciones de los teoremas de la sección 4.3.

```
function d= Teo3(P)
d=0;
if det(P)<=0
d=1;
end
n= length(P);
t=1;
for i=1:n
t=t*P(i,i);
end
if det(P)>t
d=1;
end
```

Si 1 significa que cumple las condiciones del Teorema , por lo que no existe una matriz generadora exacta.

```
function [c,d]= Teo4(P)
c=0;
d=0;
if det(P)>0.5
c=1;
end
n= length(P);
if (det(P)>0.5)&( distE(P, eye(n) )<0.5 )
d=1;
end
if ( Teo4.3(P) == 1)&(det (P)>exp( -pi ) )
d=1;
```

```
end
```

Si $c = 1$ significa que P tiene a lo más una generadora.

Si $d = 1$ significa que la única posible matriz generadora es $ln(P)$.

```
function d= Teo4.3(P)
d=1;
A=eig(P);
n=size(A);
for i=1:n
for j=(i+1):n
if A(i) == A(j)
d=0;
end
end
end
```

Si $d = 1$ entonces todos los eigenvalores son distintos.

```
function [e,f]= Teo5(P)
n= mean(size (P) );
Z= zeros( n, 1);
e=0;
f=0;
if ( imag ( eig(P) ) == Z )&( Teo4.3(P) == 1 )
if Teo5.1(P) == 1
e=1;
end
if Teo5.2(P) == 1
f=1;
end
end
```

Si $e = 1$ significa que $ln(P)$ es la única matriz real Λ tal que $\exp(\Lambda) = P$.

Si $f = 1$ significa que no existe matriz real Λ tal que $\exp(\Lambda) = P$.

```
function d= Teo5.1(P)
d=1;
A=eig(P);
n=size(A);
```

```
for i=1:n
if A(i)<=0
d=0;
end
end
```

Si $d = 1$ entonces todos los eigenvalores son estrictamente positivos.

```
function d= Teo5.2(P)
d=0;
A=eig(P);
n=size(A);
for i=1:n
if A(i)<0
d=1;
end
end
```

Si $d = 1$ entonces existe algún eigenvalor negativo.

Este programa construye la matriz de intensidad Λ a partir de una matriz de transición P por medio del método de Jarrow-Lando-Turnbull.

```
function L= JLT(P)
n=length(P);
T= zeros(n);
for i=1:(n-1)
T(i,i)= log(P(i,i))-(P(i,i)*log(P(i,i))/(P(i,i)-1));
end
L=zeros(n);
for i=1:(n-1)
for j=1:n
L(i,j)= P(i,j)*log(P(i,i))/(P(i,i)-1);
end
end
L=L+T;
```

Este programa construye la matriz de intensidad Λ a partir de una matriz de transición P por medio del método de Jarrow-Lando-Turnbull.

```
function [S,L] = IRW(P,k)
n= mean(size (P));
```

```
A=eig(P);
R=real(A);
I=imag(A);
C=(R - ones(n,1)).^ 2 + I.^ 2;
C=sort(C);
S=C(n);
L=zeros(n);
for i=1:k
L=L + ( (-1)^ (i-1))*((P- eye(n))^ i )/i ;
end
```

Este programa ajusta una matriz de intensidad Λ a partir del método AD.

```
function R = AD(L)
n= mean(size (L));
A=zeros(n);
for i=1:n
for j=1:n
if (i = j)&(L(i,j)<0)
A(i,j)=1;
L(i,j)=0;
end
end
end
T= eye(n).*L;
V=sum((L-T)');
for i=1:n
for j=1:n
if (i==j)
L(i,j)= -V(i);
end
end
end
R=L;
```

Este programa resuelve matrices de transición por programación dinámica.

```
function F = PD(P,k)
```

```
P=P^ (1/k)
n=mean(size (P));
lambda=zeros(1,n);
A=zeros(n);
for i=1:n
for j=1:n
if (P(i,j)<0 )
A(i,j)=1;
end
end
end
S=sum(A')
```

Se pone un transpuesto dentro de la suma para que sume por renglones.

```
for i=1:n
if ( S(i)>0 )
T=sum(P');
lambda(i)= (1/n)* (T(i)-1);
end
end
for i=1:n
for j=1:n
P(i,j)=P(i,j)-lambda(i);
end
end
B=zeros(n);
I=zeros(n);
D=zeros(n);
J=zeros(n);
for i=1:n
[D(i,:),J(i,:)]=sort(P(i,:));
for j=1:n
B(i,j)=D(i, n-j+1);
I(i,j)=J(i, n-j+1);
end
```

```
end
B
I
k=zeros(n,1);
C=zeros(n);
for i=1:n
if ( S(i) >0)
for j=1:n
C(i,j)=sum(B(i,:))-i*B(i,j);
end
end
end
C
D=zeros(n);
for i=1:n
for j=1:n
if (C(i,j)<=1)
D(i,j)=C(i,j);
end
end
end
D
[Y,K]=max(D');
Y
K
E=zeros(1,n);
C=B;
for i=1:n
if (S(i)>0)
for j=1:K(i)
E(i)=E(i)+B(i,j);
end
for j=1:K(i)
C(i,j)=B(i,j)+ (1/K(i))*(1-E(i));
```

```
end
for j=(K(i)+1):n
C(i,j)=0;
end
end
end
C
F=zeros(n);
for i=1:n
v=I(i,:);
for j=1:n
F(i,v(j)) = C(i,j);
end
end
```

Este programa resuelve matrices de transición por programación dinámica método OG.

```
function H = OG(P)
P=logm(P)
n=length(P);
A=zeros(n);
for i=1:n
for j=1:n
if (P(i,j)<0 )
A(i,j)=1;
end
end
end
S=sum(A’);
S=S’;
L=zeros(n,1);
for i=1:n
if ( S(i)>1 )
T=sum(P’);
L(i)= (1/n)* (T(i));
end
```

```
end
L;
for i=1:n
for j=1:n
P(i,j)=P(i,j)-L(i);
end
end
[Y,I]=sort(P');
Y=Y';
I=I';
M=cumsum(Y');
M=M';
J=fliplr(Y);
A=zeros(n);
for l=2:(n-1)
for i=1:n
for j=1:n
if ( S(i) > 1 )
A(i,l) = (n-l+1)*Y(i,l+1);
end
end
end
end
A;
F=cumsum(J');
B=fliplr(F');
G=zeros(n);
for l=2:(n-1)
for i=1:n
for j=0:(n-2)
if ( S(i)> 1 )
G(i,l) = Y(i,1)+B(i,(l+1));
end
end
```

```
end
end
G;
J=ones(n)*(n+1);
for i=1:n
for j=1:n
if ( A(i,j)> G(i,j) )
J(i,j)=j;
end
end
end
B=sort(J');
J=B';
E=zeros(n,1);
for i=1:n
if ( S(i) > 1)
E(i)=J(i,1);
end
end
E;
for i = 1:n
if ( S(i) > 1)
for k = 2:E(i)
Y(i,k)=0;
end
end
if (S(i) >1)
for k = (E(i)+1):n
Y(i,k) = Y(i,k) + (1/(n-E(i)+1))*(M(i,E(i))-M(i,1));
end
end
if (S(i) >1)
Y(i,1) = Y(i,1) +(1/(n-E(i)+1))*(M(i,E(i))-M(i,1));
end
```

```
end
H=zeros(n);
for i=1:n
v=I(i,:);
for j=1:n
H(i,v(j)) = Y(i,j);
end
end
```

Este programa calcula las metrica L1, L2 y L máx entre la matriz de transición original y la estimadas por los distintos métodos

```
function [L1,L2,LM] = M1(P,E)
n= length(P);
A=zeros(n);
for i=1:n
for j=1:n
A(i,j)=P(i,j)-E(i,j);
end
end
for i=1:n
for j=1:n
if( A(i,j) < 0 )
A(i,j)=-A(i,j);
end
end
end
S=sum(A');
A=S';
S=sum(A);
L1=S;
B=zeros(n);
for i=1:n
for j=1:n
B(i,j)=(P(i,j)-E(i,j))^ 2;
end
end
```

```
S=sum(B');
B=S';
S=sum(B);
L2=S;
C=zeros(n);
for i=1:n
for j=1:n
C(i,j)=P(i,j)-E(i,j);
end
end
for i=1:n
for j=1:n
if( C(i,j)<0 )
C(i,j)=-C(i,j);
end
end
end
M=max(C');
LM=max(M');
```

Referencias

[1] Shafer, Glenn/Vovk, Vladimir. (2001). *Probability and Finance It's Only a Game!*. John Wiley and Sons, INC., New York

[2] Merton, Robert C. (1990). *Continuous-Time Finance*. Basil Blackwell, Massachusetts

[3] Luenberger, David G. (1998). *Investment Science*. Oxford University Press, New York

[4] Haugen, Robert A. (1997). *Modern Investment Theory*. Univesity of California at Irvine

[5] Cinlar, Erhan. (1975). *Introduction to Stochastic Processes*. Prentice-Hall.

[6] Brémaud, Pierre. (1998). *Markov Chains: Gibbs fields, Monte Carlo simulation and queues*. New York, Springer-Verlag.

[7] Tudor, Constantin. (1997). *Procesos Estocásticos*. México, Sociedad Matemática Mexicana .

[8] Karatzas, Ioannis/Shreve,Steven E. (1991). *Brownian motion and stochastics calculus*. New York, Springer-Verlag.

[9] Shiryaev, Al'bert Nikolaevich. (1995). *Probability*. New York, Springer-Verlag.

[10] Ammann, Manuel. (2001). *Credit Risk Valuation: Methods, Models, and Aplications*. New York, Springer-Verlag.

[11] Bingham, Nicholas H.. (1998). *Risk-neutral valuation: pricing and hedging of financial derivatives*. London, Springer-Verlag.

[12] Shreve, Steven E.. (2003). *Stochastic calculus for finances II*. New York, Springer-Verlag.

[13] Kreinin, Alexander/Sidelnikova, Marina. (2001). *Regularization Algorithms for Transition Matrices*. Algo Research Quartely, Vol. 4, NOS.1/2, Marzo/Junio.

[14] Kijima, Massaaki. (2003). *Stochastic Processes with Aplications to Finance*. Chapman and Hall.

[15] Jarrow/Lando/Turnbull. (1997). *A Markov Model for the Term Structure of credit Risk Spreads*. The Review of Financial Studies, Vol. 10, No. 2, pág. 481-523.

[16] Israel/Rosenthal/Wei. (2001). *Finding Generators for Markov Chains via empirical transition matrices, with applications to credit ratings*. Mathematical Finance, Vol. 11, No. 2 (April 2001), pág. 245-265.

[17] Shiryaev, Al'bert Nikolaevich. (1999). *Essential of Stochastic Finance*. London, Springer-Verlag.

[18] *www.mooys.com*.

[19] *www.standarandpoors.com*.

[20] Bladt, Mogens/Sørensen, Michael. (2005). *Statistical interference for discretely observed Markov jump processes*. J.R. Statist. Soc. B(2005), No 67, Part 3, pp 395-410.

[21] Bladt, Mogens/Sørensen, Michael. (2005). *Efficient estimation of transition rates between credit ratings from observations at discrete time points*.

[22] Lando, David. (1999). *Some Elements of Rating-Based Credit Risk Modeling*.

[23] Jarrow/Lando/Yu. (2003). *Default Risk and Diversification: Theory and Aplications*.

[24] Bielecki, Tomasz R./Rutkowski, Marek. (2002). *Credit risk: modeling, valuation and hedging*. Berlin, Springer.

[25] Bluhm, Christian/Overbeck, Ludger/Wagner, Christoph. (2003). *An introduction to credit risk modeling*. Boca Raton, Florida; Chapman and Hall/CRC.

[26] Coyle, Brian. (2000). *Measuring credir risk*. Canterbury, United Kigdom; Chartered Institute of Bankers.

[27] Snell, J. (1988). *Introduction to Probability*. New York; Random House.