CIMAT

Centro de Investigación en Matemáticas, A.C.

OPTIMIZACIÓN DE PORTAFOLIOS CON ESTRATEGIAS CONVERGENTES UN ENFOQUE BAJO INFORMACIÓN PARCIAL Y COMPLETA

T E S I S

Que para obtener el grado de

Maestra en Ciencias

con Orientación en

Probabilidad y Estadística

Presenta

Jessica Mariana Villar García

Directores de Tesis:

Dr. José Luis Ángel Pérez Garmendia

Dr. Antonio Murillo Salas

Autorización de la versión final

Guanajuato, Gto. Noviembre de 2020

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología (CONACyT), por la beca que me fue otorgada al igual que al Centro de Investigación en Matemáticas A.C. (CIMAT A.C.) por todas las facilidades y apoyo para llevar a cabo mis estudios de maestría.

Agradezco a mis asesores, los Dres. José Luis Ángel Pérez Garmendia y Antonio Murillo Salas, por su orientación y apoyo durante el desarrollo de este trabajo. Finalmente, a mis sinodales, la Dra. Ekaterina Todorova Kolkovska y el Dr. Daniel Hernández Hernández, por su tiempo en leer y corregir el trabajo para su finalización.

Resumen

En este trabajo se aborda el problema de optimización de portafolios con estrategias de inversión tipo convergentes, en particular trading a pares. Tales estrategias buscan explotar divergencias (o *errores*) en el precio de un activo. Esto mediante la compra y venta de un par de activos cointegrados, es decir, activos que presentan similitud en el histórico de sus precios. Se supone además que los errores en los precios asociados a los activos cointegrados están regidos por una cadena de Markov. De aquí surgen dos enfoques: completamente observable, donde la dinámica de la cadena de Markov es conocida por el inversionista; y parcialmente observable, esto es, la cadena se observa a través de un filtro.

Para resolver este problema se utilizaron técnicas de control estocástico: a partir de la ecuación de Hamilton-Jacobi-Bellman (HJB) se obtiene la estrategia óptima para luego proponer un candidato a la función de valor del problema de control asociado. Finalmente se verifica que, en efecto, tal candidato coincide con la solución de la ecuación HJB. Para el caso parcialmente observable se resuelve primero el problema de filtros, el cual permite pasar del problema de información parcial a uno equivalente de información completa.

Palabras clave: Control estocástico, información parcial, filtros, optimización de portafolios, trading convergente.

Índice general

Introducción

La optimización de portafolios ha sido un tema clásico y de gran interés dentro del área de finanzas y desde un punto de vista probabilista, en particular de la teoría de procesos estocásticos [12]. A grandes rasgos, el problema consiste en encontrar la estrategia óptima que maximice la utilidad esperada de un portafolio de inversión [10]. Los enfoques desde los que se ha estudiado este problema han sido variados y se han ido adaptando según las necesidades de los inversionistas, un ejemplo de esto son las estrategias de inversión de tipo convergente [2].

La comercialización convergente busca explotar divergencias temporales en el precio de un activo respecto al precio justo. La estrategia típica consiste en, simultáneamente comprar activos subestimados y vender en corto aquellos que estén sobrevalorados, con la expectativa de que en alguna fecha futura tales divergencias, o *errores*, se habrán ajustado, beneficiándose así de la convergencia de los precios. Dentro de este ámbito una de las estrategias tradicionales es el *trading a pares*, éstas consisten en la compra y venta de un par de activos cointegrados, esto es, activos que presentan similitud en el histórico de sus precios.

En este trabajo se busca resolver el problema de optimización maximizando la utilidad esperada de portafolios con estrategias de inversión tipo convergentes, en particular *trading a pares*. El desarrollo estará basado en el artículo *Optimal convergencce trading with unobservable pricing errors* [2] en donde se extiende el modelo implementado en [8]. El objetivo en específico es resolver el problema desde dos perspectivas: en la primera se supone que los *errores* en los precios asociados a los activos cointegrados están regidos por una cadena de Markov a tiempo continuo y con espacio de estados discreto, la cual no puede ser observada por el inversionista. La razón detrás de este supuesto se debe a que tales *errores* dependen de factores económicos o de mercado, que por su naturaleza son estocásticos y por ende resulta natural suponer que el inversionista no tenga acceso a este tipo de información. A este problema se le conoce como *información parcial*; la segunda será bajo *información completa*, esto es, bajo el supuesto de que se tiene acceso a tal información.

Aunado a lo anterior, bajo las perspectivas antes mencionadas, se analizarán carteras de inversión *beta neutral*, es decir, que permitan eliminar por completo el riesgo de mercado. Este tipo de carteras suelen ser utilizadas por inversionistas que implementan estrategias tipo *trading a pares* [2], por ello la relevancia de estudiar este enfoque.

La estructura de este trabajo será como sigue. En el Capítulo 1 se presentarán una serie de herramientas de cálculo estocástico, así como una breve introducción a la teoría de control estocástico y los resultados más relevantes de este campo. Esto nos permitirá introducir el problema de optimización de portafolios y su relación con optimización estocástica, lo cual se abordará en el Capítulo 2. Una vez introducidas tales herramientas, se describirá el modelo bajo el cual se planteará el problema de control asociado a los casos de información completa e información parcial. Así, podremos abordar el enfoque de información completa en el Capítulo 3, el resultado principal será obtener la estrategia óptima y caracterizar la función de valor asociada al problema de control. Para el caso de información parcial primero se resolverá el problema de filtros. Ello nos permitirá pasar del problema de información parcial a uno equivalente de información completa. Lo anterior se expondrá en el Capítulo 4.

Capítulo 1

Preliminares

1.1. Cálculo estocástico

Esta sección estará basada en [9], la finalidad es dar un somero repaso de algunas herramientas de cálculo estocástico que serán necesarias durante el desarrollo del trabajo.

Por brevedad, definiremos a la integral estocástica con respecto a semimartingalas continuas. Sin embargo, por ser de utilidad para el Capítulo 3, enunciaremos la fórmula de Itô para el caso de semimartingalas generales. Se omitirá la demostración de los resultados al quedar fuera de los objetivos de la tesis.

Definiciones y notación

Las semimartingalas se pueden representar como la suma de una martingala local y un proceso de variación finita. En este apartado introduciremos algunas definiciones y notación relacionadas con tales conceptos.

Consideraremos $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ un espacio de probabilidad filtrado, donde la filtración $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ satisface las condiciones usuales, es decir, $\mathbb{F}$ es $\mathbb{P}$-completa y continua por la derecha.

Definición 1.1.1. *Sea $T \geq 0$. Una función continua $a : [0,T] \to \mathbb{R}$, tal que $a(0) = 0$ se dice que es de variación finita si existe una medida con signo μ en $[0,T]$ tal que $a(t) = \mu([0,T])$, para toda $t \in [0,T]$.*

Denotaremos por $|\mu| = \mu_+ + \mu_-$ a la variación total de μ. Para $f : [0,T] \to \mathbb{R}$ medible y tal que

$\int_{[0,T]} |f(s)||\mu|(ds) < \infty$ *se tiene que*

$$\int_0^T f(s)da(s) = \int_{[0,T]} f(s)\mu(ds).$$

Definición 1.1.2. *Un proceso $A = (A_t)_{t\geq 0}$ es de variación finita si es un proceso adaptado y tal que sus trayectorias son funciones de variación finita, con probabilidad uno. Si además las trayectorias son no decrecientes, entonces se dice que A es un proceso creciente.*

Definición 1.1.3. *Un proceso adaptado $M = (M_t)_{t\geq 0}$ con trayectorias continuas y tal que $M_0 = 0$ c.s. se dice que es una martingala local si existe una sucesión no decreciente de tiempos de paro $(T_n, n \in \mathbb{N})$ tal que $T_n \uparrow \infty$ y para toda n el proceso parado $M_{T_n} := \{M_{T_n \wedge t} : t \geq 0\}$ es una martingala uniformemente integrable.*

Más generalmente, si $M_0 \neq 0$, se dice que M es una martingala local si $M_t = M_0 + N_t$, donde el proceso N es una martingala local que inicia en cero.

Dentro de la teoría de cálculo estocástico la variación cuadrática juega un rol muy importante, concepto que se define como sigue:

Teorema 1.1.1. *Sea $M = (M_t)_{t\geq 0}$ una martingala local continua. Existe un único proceso creciente, denotado por $([M,M]_t)_{t\geq 0}$, tal que $M_t^2 - [M,M]_t$ es una martingala local continua.*

Además, para $t > 0$ fijo, si $\Pi_n = \{t_0^n, t_1^n, \ldots, t_{p_n}^n\}$, con $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$, es una sucesión no decreciente de particiones del intervalo $[0,t]$ con norma $\|\Pi_n\| = \max\limits_{1\leq k\leq p_n} |t_k^n - t_{k-1}^n|$ tendiendo a cero, se tiene que

$$[M,M]_t = \lim_{n\to\infty} \sum_{k=1}^{p_n} |M_{t_k^n} - M_{t_{k-1}^n}|^2$$

en probabilidad. Al proceso $[M,M]$ se le llama variación cuadrática de M.

Demostración. Ver Le Gall [9, Teorema 4.9]. ■

Definición 1.1.4. *Sean M y N martingalas locales. Definimos a la* ***variación cruzada*** *como*

$$[M,N]_t := \frac{1}{2}\big([M+N, M+N]_t - [N,N]_t - [M,M]_t\big).$$

La siguiente proposición concluye propiedades similares a las del teorema anterior, respecto de la variación cruzada

Proposición 1.1.1. *Sean M y N martingalas locales. Entonces,*

- $[M,N]_t$ *es el único proceso de variación finita tal que* $M_tN_t - [M,N]_t$ *es martingala local continua.*

- *Si* $\Pi_n = \{t_0^n, t_1^n, \ldots, t_{p_n}^n\}$, *con* $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$, *es una sucesión no decreciente de particiones del intervalo* $[0,t]$ *con norma* $\|\Pi_n\| = \max\limits_{1 \leq k \leq p_n} |t_k^n - t_{k-1}^n|$ *tendiendo a cero, se tiene que*

$$[M,N]_t = \lim_{n\to\infty} \sum_{k=1}^{p_n} (M_{t_k^n} - M_{t_{k-1}^n})(N_{t_k^n} - N_{t_{k-1}^n}),$$

en probabilidad.

Demostración. Ver Le Gall [9, Proposición 4.15] ∎

Observación 1.1.1. *Se puede demostrar que, si* $W = (W^1, \ldots, W^d)$ *es un movimiento browniano con valores en* $\mathbb{R}^d$, *entonces* $[W^k, W^j]_t = \delta_{kj} t$, $1 \leq k, j \leq d$ *(ver Le Gall [9] pág.117).*

Con las definiciones anteriores podemos ahora introducir a las semimartingalas.

Definición 1.1.5. *Un proceso estocástico* $X = (X_t)_{t\geq 0}$ *es una semimartingala continua si se puede escribir como* $X_t = X_0 + M_t + A_t$, *donde* M *es una martingala local con* $M_0 = 0$ *y* A *es un proceso de variación finita.*

Notación 1.1.1. *Para el desarrollo de la siguiente sección se utilizará la siguiente notación:*

- H^2: *espacio de martingalas continuas acotadas en* L^2, *esto es,* $\sup\limits_n \mathbb{E}(M_n^2) < +\infty$, *y tales que* $M_0 = 0$.

- *Para* $M \in H^2$, *denotaremos por* $L^2(M)$ *al conjunto de procesos progresivamente medibles* H *tales que*

$$\mathbb{E}\left(\int_0^t H_s^2 d[M,M]_s\right) < \infty.$$

Integral estocástica

El calculo estocástico es una de las herramientas centrales en la teoría de matemáticas financieras. Esta rama surge de la necesidad de darle un significado a las ecuaciones diferenciales estocásticas [14]. En este aparatado se introduce el concepto de integral estocástica, la idea será definirla para martingalas, luego para martingalas locales y finalmente para semimartingalas.

Definición 1.1.6. *Se dice que un proceso es simple si es progresivamente medible y tiene la siguiente representación*

$$H_s(w) = \sum_{i=0}^{p-1} H_i(\omega)\mathbb{1}_{(t_i,t_{i+1}]}(s),$$

donde $0 = t_0 < \cdots < t_p$ *y* $H_i \in \mathcal{F}_{t_i}$, $i \in \{0, 1, \ldots, p-1\}$.

Denotaremos por $\mathcal{S}$ *al conjunto de procesos simples.*

Teorema 1.1.2. *Sea* $M \in H^2$ *y* $H \in \mathcal{S}$. *El operador* $H \cdot M$, *dado por*

$$(H \cdot M)_t = \sum_{i=0}^{p-1} H_i(M_{t_{i+1}\wedge t} - M_{t_i \wedge t}),$$

define a un proceso en H^2. *Además, se satisface lo siguiente:*

1. *El mapeo* $H \to H \cdot M$ *induce una isometría de* $L^2(M)$ *a* H^2.
2. $H \cdot M$ *es la única martingala en* H^2 *tal que* $[H \cdot M, N] = H \cdot [M, N]$, *para toda* $N \in H^2$.
3. *Si* T *es un tiempo de paro entonces,* $(\mathbb{1}_{[0,T]}H) \cdot M = (H \cdot M)^T = H \cdot M_{\cdot\wedge T}$.

Utilizaremos la siguiente notación:

$$(H \cdot M)_t = \int_0^t H_s dM_s$$

y diremos que $H \cdot M$ *es la integral estocástica de* H *con respecto de* M.

Demostración. Ver Le Gall [9, Proposición 5.4]. ■

Observación 1.1.2. *Del punto 2 del teorema anterior podemos decir que la integral estocástica conmuta con la variación cuadrática. En particular, si* $M \in H^2$ *y* $H \in L^2(M)$ *entonces,* $[H \cdot M, H \cdot M] = H^2 \cdot [M, M]$.

El siguiente resultado nos permitirá concluir ciertas propiedades respecto a los momentos de la integral estocástica.

Proposición 1.1.2. *Sea* $H \in L^2(M)$. *Si* K *es un proceso progresivamente medible, entonces* $KH \in L^2(M)$ *si y sólo si* $K \in L^2(H \cdot M)$. *Si lo anterior se satisface entonces,* $(KH) \cdot M = K \cdot (H \cdot M)$.

Demostración. Ver Le Gall [9, Proposición 5.5]. ■

Observación 1.1.3. *De la observación y proposición anteriores se sigue que: si $M \in H^2$, $N \in H^2$, $K \in L^2(N)$ y $H \in L^2(M)$, como $H \cdot M$ y $K \cdot N$ son martingalas en H^2, entonces*

$$\mathbb{E}\bigg(\int_0^t H_s dM_s\bigg) = 0 \quad y \quad \mathbb{E}\bigg(\bigg(\int_0^t H_s dM_s\bigg)\bigg(\int_0^t K_s dN_s\bigg)\bigg) = \mathbb{E}\bigg(\int_0^t H_s K_s d[M,N]_s\bigg), \quad t \in [0,\infty],$$

en particular

$$\mathbb{E}\bigg(\bigg(\int_0^t H_s dM_s\bigg)^2\bigg) = \mathbb{E}\bigg(\int_0^t H_s^2 d[M,M]_s\bigg).$$

En lo siguiente extenderemos la definición de la integral estocástica para martingalas locales.

Para M una martingala local continua, denotaremos por $L^2_{loc}(M)$ al conjunto de procesos H progresivamente medibles tales que

$$\int_0^t H_s^2 d[M,M]_s < \infty, \quad \forall t \geq 0, \quad \text{c.s.}$$

Teorema 1.1.3. *Sea M una martingala local continua. Para todo $H \in L^2_{loc}(M)$, existe una única martingala local continua que inicia en cero, denotada por $H \cdot M$, tal que $[H \cdot M, N] = H \cdot [M,N]$, para toda martingala local continua N. Además, se satisface lo siguiente*

1. *Si T es un tiempo de paro, se tiene que $(\mathbb{1}_{[0,T]}H) \cdot M = (H \cdot M)^T = H \cdot M_{\cdot \wedge T}$.*
2. *Si $H \in L^2_{loc}(M)$ y K es un proceso progresivamente medible entonces, $K \in L^2_{loc}(H \cdot M)$ si y sólo si $HK \in L^2_{loc}(M)$. Si lo anterior se satisface entonces, $H \cdot (K \cdot M) = HK \cdot M$.*
3. *Si $M \in H^2$ y $H \in L^2(M)$, la definición de $H \cdot M$ es consistente con la del Teorema 1.1.2.*

Demostración. Ver Le Gall [9, Teorema 5.6]. ■

Observación 1.1.4. *Sea M una martingala local continua y $H \in L^2_{loc}(M)$. Bajo la condición*

$$\mathbb{E}\bigg(\int_0^t H_s^2 d[M,M]_s\bigg) < \infty,$$

las conclusiones en las Observaciones 1.1.2 y 1.1.3 se preservan.

Integral estocástica para semimartingalas

Finalmente podemos extender la definición de integral estocástica para el caso de semimartingalas continuas.

Diremos que un proceso H progresivamente medible es localmente acotado si, para toda $t \geq 0$, $\sup_{s \leq t} |H_s| < \infty$ c.s. Si H es progresivamente medible y localmente acotado entonces, para todo proceso de variación finita V se satisface casi seguramente, para toda $t \geq 0$,

$$\int_0^t |H_s||dV_s| < \infty.$$

Definición 1.1.7. *Sea X una semimartingala continua con descomposición $X = X_0 + M + V$. Si H es un proceso progresivamente medible y localmente acotado, la integral estocástica $H \cdot X$ es la semimartingala continua con descomposición $H \cdot X = H \cdot M + H \cdot V$ y la denotamos como*

$$H \cdot X = \int_0^t H_s dX_s.$$

Proposición 1.1.3. *Sean X y H como en la definición anterior, la integral estocástica $H \cdot X$ satisface*

1. *Si T es un tiempo de paro $(H \cdot X)^T = H \mathbb{1}_{[0,T]} \cdot X = H \cdot X_{\cdot \wedge T}$.*
2. *Si H es de la forma*

$$H_s(w) = \sum_{i=0}^{p-1} H_i(\omega) \mathbb{1}_{(t_i, t_{i+1}]}(s),$$

donde $0 = t_0 < \cdots < t_p$ y $H_i \in \mathcal{F}_{t_i}$, $i \in \{0, 1, \ldots, p-1\}$, entonces

$$(H \cdot X)_t = \sum_{i=0}^{p-1} H_i (X_{t_{i+1} \wedge t} - X_{t_i \wedge t}).$$

3. *Si H y K son procesos progresivamente medibles y localmente acotados, entonces $H \cdot (K \cdot X) = (HK) \cdot X$.*

Demostración. Ver Le Gall [9] pág. 109. ∎

El siguiente teorema nos permite extender la definición de integral estocástica para integrandos continuos.

Proposición 1.1.4. *Sea X una semimartingala continua y H un proceso adaptado con trayectorias continuas. Entonces, para cada $t > 0$ y $\Pi_n = \{t_0^n, t_1^n, \ldots, t_{p_n}^n\}$, con $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$, sucesión de particiones del intervalo $[0, t]$ con norma $\|\Pi_n\| = \max_{1 \leq k \leq p_n} |t_k^n - t_{k-1}^n|$ tendiendo a cero, tenemos que*

$$\lim_{n \to \infty} \sum_{i=0}^{p_n - 1} H_{t_i^n} (X_{t_{i+1}^n} - X_{t_i}^n) = \int_0^t H_s dX_s,$$

en probabilidad.

Demostración. Ver Le Gall [9, Proposición 5.9]. ■

Fórmula de Itô

La fórmula de Itô es uno de los resultados más importantes en cálculo estocástico. Tal resultado concluye que si aplicamos una función continua y dos veces diferenciable a un vector de semimartingalas, entonces el proceso que resulta de tal aplicación es también una semimartingala y además se tiene una forma explícita de su descomposición.

Como mencionamos anteriormente, enunciaremos la fórmula de Itô para el caso de semimartingalas generales ya que tal resultado será útil en el Capítulo 3.

Teorema 1.1.4. *(Fórmula de Itô) Sea $X = (X^1, \ldots, X^n)$ un vector de semimartingalas y $f : \mathbb{R}^n \to \mathbb{R}$ una función de clase C^2. Entonces, $f(X)$ es una semimartingala y se satisface que*

$$f(X_t) - f(X_0) = \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}(X_{s^-}) dX_s^i + \frac{1}{2} \sum_{1 \leq i,j \leq n} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_{s^-}) d[X^i, X^j]_s^c$$
$$+ \sum_{0 \leq s \leq t} \left(f(X_s) - f(X_{s^-}) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_{s^-}) \Delta X_s^i \right),$$

donde $X_{s^-} = \lim\limits_{s \to t, s<t} X_s$ y $[X,X]^c$ denota la parte continua, trayectoria por trayectoria, de $[X,X]$.

Demostración. Ver Protter [13, Teorema 2.7.22]. ■

Para semimartingalas continuas lo expondremos como un corolario del caso anterior.

Corolario 1.1.1. *Sea $X = (X^1, \ldots, X^n)$ un vector de semimartingalas continuas y $f : \mathbb{R}^n \to \mathbb{R}$ una función de clase C^2. Entonces $f(X)$ es una semimartingala y se satisface que*

$$f(X_t) - f(X_0) = \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}(X_s) dX_s^i + \frac{1}{2} \sum_{1 \leq i,j \leq n} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) d[X^i, X^j]_s.$$

Demostración. Ver Le Gall [9, Teorema 5.10]. ■

Observación 1.1.5. *La **fórmula del producto** o **integración por partes** es un caso particular de la fórmula de Itô, que se obtiene al considerar $n = 2$ y $f(x,y) = xy$. Si X y Y son semimartingalas continuas, entonces*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_t dY_t + \int_0^t Y_t dX_t + [X,Y]_t.$$

Teorema de caracterización de Lévy del movimiento browniano

Si W es un movimiento browniano con valores en $\mathbb{R}^d$, en la Observación 1.1.1 vimos que $[W^k, W^j] = \delta_{kj}t$, $1 \leq k, j \leq d$. El teorema de caracterización de Lévy concluye que esta propiedad caracteriza al movimiento browniano entre las martingalas locales continuas. Más precisamente,

Teorema 1.1.5. *Sea $X = (X^1, \ldots, X^d)$ un proceso adaptado con trayectorias continuas. Las siguientes son equivalentes:*

1. *X es un $\mathbb{F}$ movimiento browniano.*
2. *El proceso X es un vector de martingalas locales continuas y $[X^i, X^j] = \delta_{ij}t$, $i, j \in \{1, \ldots, d\}$.*

En particular, una martingala local continua M es un $\mathbb{F}$ movimiento browniano si y sólo si $[M, M] = t$, para $t \geq 0$, o equivalentemente, si y sólo si $M_t^2 - t$ es una martingala local continua.

Demostración. Ver Le Gall [9, Teorema 5.12]. ∎

Ecuaciones diferenciales estocásticas

En este apartado definiremos el concepto de ecuación diferencial estocástica y las condiciones necesarias para la existencia y unicidad de soluciones.

Definición 1.1.8. *Sean σ y b funciones medibles localmente acotadas definidas en $\mathbb{R}^+ \times \mathbb{R}^d$ con valores en $M_{d\times m}(\mathbb{R})$ y $\mathbb{R}^d$ respectivamente, donde $M_{d\times m}(\mathbb{R})$ denota el conjunto de matrices $d \times m$ con coeficientes reales. Denotamos por $b(t, x) = (b_i(t, x))_{1\leq i\leq d}$ y $\sigma(t, x) = (\sigma_{ij}(t, x))_{1\leq i\leq d, 1\leq j\leq m}$.*

De manera muy somera, una solución de la ecuación diferencial estocástica

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

consiste en:

1. *un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, donde la filtración siempre se supone que es completa,*
2. *un $(\mathcal{F}_t)$-movimiento browniano m dimensional $W = (W^1, \ldots, W^m)$, que inicia en cero,*
3. *un proceso $X = (X^1, \ldots, X^d)$ adaptado a la filtración $(\mathcal{F}_t)$, con valores en $\mathbb{R}^d$ y trayectorias continuas tal que*

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \quad t \in [0, \infty),$$

es decir, para cada $i \in \{1, \ldots, d\}$,

$$X_t^i = X_0^i + \int_0^t b_i(s, X_s)ds + \sum_{j=1}^m \int_0^t \sigma_{ij}(s, X_s)dW_s^j.$$

Si además se cumple que $X_0 = x \in \mathbb{R}^d$, decimos que X es solución de $E_x(\sigma, b)$.

Se tienen distintas nociones de soluciones y unicidad para una ecuación diferencial estocástica $E(\sigma, b)$:

Definición 1.1.9. *Para la ecuación diferencial estocástica $E(\sigma, b)$ decimos que*

- *existe una solución débil si, para cada $x \in \mathbb{R}^d$ existe una solución de $E_x(\sigma, b)$,*
- *existe una solución débil y unicidad débil si además para cada $x \in \mathbb{R}^d$, todas las soluciones de $E_x(\sigma, b)$ tienen la misma ley,*
- *unicidad por trayectorias si, al fijar el espacio de probabilidad filtrado $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ y el $(\mathcal{F}_t)$-movimiento browniano W, dos soluciones X y X' tales que $X_0 = X'_0$, son indistinguibles c.s.*

Más aún, decimos que X es solución fuerte de $E_x(\sigma, b)$ si es adaptado respecto a la filtración canónica completa de W.

Teorema 1.1.6. *Supongamos que las funciones b y σ son continuas y satisfacen*

$$\|b(t,x) - b(t,y)\| \leq K\|x - y\|, \qquad \|\sigma(t,x) - \sigma(t,y)\| \leq K\|x - y\|,$$

para cada $x, y \in \mathbb{R}^d$ y $t \in [0, \infty)$, donde K es una constante positiva.

Entonces existe una solución única por trayectorias de $E(\sigma, b)$ y, para cada elección de un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ y un $(\mathcal{F}_t)$-movimiento browniano W, para cada $x \in \mathbb{R}^d$, existe una única solución fuerte continua de $E_x(\sigma, b)$. Además, si $X_0 = \xi$, para $\xi \in \mathbb{R}^d$ independiente de W y con segundo momento finito, para cada $T > 0$ existe una constante $C(K, T)$, tal que

$$\mathbb{E}\left(\sup_{t \leq T} \|X_t\|^2\right) \leq C(1 + \mathbb{E}(\|X_0\|^2))e^{Ct}, \quad t \in [0, T].$$

Demostración. Ver Le Gall [9, Teorema 8.3]. ∎

Observación 1.1.6. *Si las funciones σ y b fueran localmente Lipschitz es necesario contar con una condición de crecimiento lineal de la forma*

$$\|\sigma(t,x)\| \leq K(1 + \|x\|), \qquad \|b(t,x)\| \leq K(1 + \|x\|).$$

Existe una gama de ecuaciones diferenciales estocásticas de las cuales se tienen fórmulas bastante explícitas de las soluciones. A continuación se enuncia una en particular que resulta útil para los Capítulos 3 y 4.

Lema 1.1.1. *La solución a la ecuación diferencial estocástica*

$$\begin{cases} dX = (c(t) + d(t)X)dt + \sum_{l=1}^{m}(e^l(t) + f^l(t)X)dW_t^l \\ X(0) = X_0, \end{cases}$$

es

$$X(t) = \Phi(t)\Bigg(X_0 + \int_0^t \Phi^{-1}(s)\Big[c(s) - \sum_{l=1}^{m} e^l(s)f^l(s)\Big]ds\Bigg) + \int_0^t \sum_{l=1}^{m} \Phi^{-1}(s)e^l(s)dW_s^l,$$

donde,

$$\Phi(t) := \exp\Bigg(\int_0^t d(s) - \sum_{l=1}^{m} \frac{(f^l(s))^2}{2}ds + \int_0^t \sum_{l=1}^{m} f^l(s)dW_s^l\Bigg).$$

Demostración. Ver Evans [4, Teorema 5.4.3]. ■

1.2. Control estocástico

En el campo de control estocástico se estudian sistemas dinámicos sujetos a perturbaciones aleatorias, los cuales pueden ser controlados para optimizar cierto criterio. Tal teoría se ha aplicado en distintos campos, particularmente en problemas de toma decisiones en economía y finanzas. En este apartado se presentarán los resultados principales de control estocástico. Para el desarrollo de éste se utilizarán como referencias [15] y [12].

1.2.1. Problema de optimización estocástica

Para poder formular un problema de optimización estocástica, en términos generales, se consideran los siguientes elementos:

Sea $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ un espacio de probabilidad filtrado, donde $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ satisface las condiciones usuales. Además, sea $W = \{W_t, t \geq 0\}$ un movimiento browniano definido en $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ con valores en $\mathbb{R}^d$. Denotaremos por $\mathbf{S} := [0, T) \times \mathbb{R}^d$, con $T \in [0, \infty)$.

Proceso de estados. El proceso de estados representará las variables cuantitativas a considerar para poder describir el problema de interés.

Proceso de control. Dado un subconjunto U de $\mathbb{R}^k$, denotaremos por $\mathcal{U}$ al conjunto de los procesos progresivamente medibles $\nu = \{\nu_t, t < T\}$ con valores en U. A los elementos en $\mathcal{U}$ les llamaremos *procesos de control*. Serán éstos los que influirán en la dinámica de un proceso de estados.

Proceso controlado. Consideremos $b : (t, x, u) \in \mathbf{S} \times U \to b(t, x, u) \in \mathbb{R}^d$ y $\sigma : (t, x, u) \in \mathbf{S} \times U \to \sigma(t, x, u) \in M_{n\times d}(\mathbb{R})$, funciones continuas tales que satisfacen

$$\begin{aligned} \|b(t,x,u) - b(t,y,u)\| + \|\sigma(t,x,u) - \sigma(t,y,u)\| &\leq K\|x-y\|, \\ \|b(t,x,u)\| + \|\sigma(t,x,u)\| &\leq K(1 + \|x\| + \|u\|), \end{aligned}$$

para alguna constante K independiente de (t, x, u).

Para cada proceso de control $\nu \in \mathcal{U}$, consideremos la ecuación diferencial estocástica controlada

$$dX_t = b(t, X_t, \nu_t) + \sigma(t, X_t, \nu_t)dW_t, \tag{1.1}$$

si la ecuación anterior tiene una única solución, para un valor inicial dado, entonces llamaremos al proceso X *proceso controlado*. Típicamente suelen considerarse procesos de control cuadrado integrables. Así, al ser b y σ uniformemente Lipschitz, se puede asegurar la existencia de un proceso controlado en el intervalo $[0, T']$, con $T' \leq T$.

El siguiente resultado y su demostración se puede consultar en [15].

Teorema 1.2.1. *Sea $\nu \in \mathcal{U}_0$ un proceso de control y ξ una variable aleatoria $\mathcal{F}_0$ medible y con segundo momento finito. Entonces, existe un único proceso, X^ν, $\mathbb{F}$ adaptado que satisface* (1.1), *con la condición inicial $X_0^\nu = \xi$. Además, para cada $T > 0$, existe una constante $C > 0$ tal que*

$$\mathbb{E}\left(\sup_{0\leq s\leq t} \|X_s^\nu\|^2\right) \leq C(1 + \mathbb{E}(\|\xi\|^2))e^{Ct}, \quad t \in [0, T),$$

donde $\mathcal{U}_0$ es la clase de procesos controlados progresivamente medibles y cuadrado integrables.

Los últimos dos elementos a considerar para formular el problema de optimización estocástica son la *función de costo* y los *controles admisibles*, los cuales se describen a continuación:

Función de costo. Para introducir la función de costo necesitamos los siguientes componentes: sean

$$\begin{aligned} f, k &: [0, T) \times \mathbb{R}^d \times U \to \mathbb{R}, \\ g &: \mathbb{R}^d \to \mathbb{R}, \end{aligned}$$

funciones que supondremos continuas y donde k satisface $\|k^-\|_\infty < \infty$. Además supondremos que las funciones f y g cumplen

$$|f(t,x,u)| + |g(x)| \leq K(1 + \|u\| + \|x\|^2),$$

para alguna constante K independiente de (t,x,u).

Definimos a la función de costo, J, como

$$J(t,x,\nu) := \mathbb{E}\bigg(\int_t^T \beta^\nu(t,s) f(s, X_s^{t,x,\nu}, \nu_s)ds + \beta^\nu(t,T) g(X_T^{t,x,\nu}) \mathbb{1}_{\{T<\infty\}}\bigg),$$

donde $\beta^\nu(t,s) := \exp\big(-\int_t^T k(r, X_r^{t,x,\nu}, \nu_r)dr\big)$, y $\{X_s^{t,x,\nu}, s \geq t\}$ es la solución a (1.1) con proceso de control ν y condición inicial $X_t^{t,x,\nu} = x$.

En lo anterior, la función f se puede interpretar como una función de ganancia corriente, a g como la ganancia terminal y β como un factor de descuento.

Proceso de control admisible. Para el caso en que $T < \infty$, al tener f y g crecimiento cuadrático y por ser k^- acotada, se tiene que $J(t,x,\nu)$ está bien definida para cualquier $\nu \in \mathcal{U}_0$. Para este caso, $\mathcal{U}_0$ representará el conjunto de controles admisibles.

Para el caso en que $T = \infty$, definimos al conjunto de controles admisibles como

$$\mathcal{U}_0 := \bigg\{\nu \in \mathcal{U} : \mathbb{E}\bigg(\int_0^\infty \beta^\nu(t,s)(1 + \|X_s^{t,x,\nu}\|^2 + \|\nu_s\|)ds\bigg) < \infty, \forall x\bigg\}.$$

El problema de control estocástico. Consideremos el problema de optimización

$$V(t,x) := \sup_{\nu \in \mathcal{U}_0} J(t,x,\nu), \quad \text{para } (t,x) \in \mathbf{S}.$$

El objetivo principal en un problema de optimización estocástica es encontrar el proceso de control ν que maximice la función de costo, $J(\cdot)$.

Diremos que ν^* es un control óptimo si $V(t,x) = J(t,x,\nu^*)$. Además, si el control es de la forma $\nu_s = a(s, X_s)$ para alguna función medible $a : [0,T] \times \mathbb{R}^d \to U$, entonces le llamaremos control Markoviano.

1.2.2. Principio de programación dinámica

El principio de programación dinámica juega un rol crucial dentro de la teoría de control estocástico. En términos generales, proporciona una manera conveniente de manejar un problema de optimización global resolviendo recursivamente una serie de problemas de optimización locales.

Antes de enunciar el principio de programación dinámica introducimos la siguiente notación:

Notación 1.2.1.

$$\begin{aligned} V^*(t,x) &:= \limsup_{(t',x')\to(t,x)} V(t',x'), \\ V_*(t,x) &:= \liminf_{(t',x')\to(t,x)} V(t',x'), \\ \mathcal{U}_t &:= \{\nu \in \mathcal{U}_0 : \nu \textit{ es independiente de } \mathcal{F}_t\}. \end{aligned}$$

El siguiente teorema y su demostración se puede consultar en Touzi [15, Teorema 3.3]

Teorema 1.2.2. *(Principio de programación dinámica.) Supongamos que la función de valor* V *es localmente acotada. Sea* $(t,x) \in \boldsymbol{S}$ *fijo y sea* $\{\theta^\nu, \nu \in \mathcal{U}_t\}$ *una familia de tiempos de paro independientes de* $\mathcal{F}_t$ *con valores en* $[t,T]$*. Entonces*

$$V(t,x) \leq \sup_{\nu\in\mathcal{U}_t} \mathbb{E}\bigg(\int_t^{\theta^\nu} \beta^\nu(t,s) f(s, X_s^{t,x,\nu}, \nu_s) ds + \beta^\nu(t,\theta^\nu) V^*(\theta^\nu, X_{\theta^\nu}^{t,x,\nu})\bigg).$$

Si suponemos además que la función g *es semicontinua por abajo y* $X_s^{t,x,\nu}\mathbb{1}_{[t,\theta^\nu]}(s) \in \mathbb{L}^\infty$ *para todo* $\nu \in \mathcal{U}_t$*. Entonces*

$$V(t,x) \geq \sup_{\nu\in\mathcal{U}_t} \mathbb{E}\bigg(\int_t^{\theta^\nu} \beta^\nu(t,s) f(s, X_s^{t,x,\nu}, \nu_s) ds + \beta^\nu(t,\theta^\nu) V_*(\theta^\nu, X_{\theta^\nu}^{t,x,\nu})\bigg).$$

Observación 1.2.1. *Si la función de valor* V *es continua, entonces* $V = V* = V_*$ *y así el principio de programación dinámica se reduce a*

$$V(t,x) = \sup_{\nu\in\mathcal{U}_t} \mathbb{E}\bigg(\int_t^{\theta^\nu} \beta^\nu(t,s) f(s, X_s^{t,x,\nu}, \nu_s) ds + \beta^\nu(t,\theta^\nu) V(\theta^\nu, X_{\theta^\nu}^{t,x,\nu})\bigg).$$

1.2.3. Ecuación de Hamilton-Jacobi-Bellman

La ecuación de Hamilton-Jacobi-Bellman, o también llamada la ecuación de programación dinámica, es la versión infinitesimal del Teorema 1.2.2.

Consideremos $\mathcal{S}_d$ al conjunto de matrices simétricas $d \times d$, con coeficientes reales. Definimos a $H : \mathbf{S} \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}_d$, como sigue

$$H(t,x,r,p,\gamma) := \sup_{u\in U} \big\{ -k(t,x,u)r + b(t,x,u)\cdot p + \frac{1}{2} Tr[\sigma\sigma^T(t,x,u)\gamma] + f(t,x,u) \big\},$$

y al operador lineal de segundo orden como

$$\mathcal{L}^u\phi(t,x) \coloneqq -k(t,x,u)\phi(t,x) + b(t,x,u)\cdot D\phi(t,x) + \frac{1}{2}Tr[\sigma\sigma^T(t,x,u)D^2\phi(t,x)],$$

donde D y D^2 representan el gradiente y la matriz Hessiana de ϕ, respectivamente.

Observación 1.2.2. *Para $s \geq t$ y $\phi \in \mathcal{C}^{1,2}([t,s],\mathbb{R}^d)$, por la fórmula de Itô se satisface que*

$$\begin{aligned} d\big(\beta^\nu(0,s)\phi(s,X_s^\nu)\big) &= \beta^\nu(0,s)d\phi(s,X_s^\nu) + \phi(s,X_s^\nu)d\beta^\nu(0,s) + d[\beta^\nu,\phi]_s \\ &= \beta^\nu(0,s)\big(b(s,X_s^\nu,\nu_s)\cdot D\phi(s,X_s^\nu) + \frac{1}{2}Tr[\sigma\sigma^T(s,X_s^\nu,\nu_s)D^2\phi(s,X_s^\nu)]\big)dt \\ &\quad + \beta^\nu(0,s)\big(\partial_t\phi(s,X_s^\nu)dt + D\phi(s,X_s^\nu)\cdot\sigma(s,X_s^\nu,\nu_s)dW_s\big). \end{aligned}$$

De forma integral obtenemos,

$$\begin{aligned} \beta^\nu(0,s)\phi(s,X_s^\nu) - \beta^\nu(0,t)\phi(t,X_t^\nu) = &\int_t^s \beta^\nu(0,r)\big(\partial_t + \mathcal{L}^\nu\big)\phi(r,X_r^\nu)dr \\ &+ \int_t^s \beta^\nu(0,r)D\phi(r,X_r^\nu)\cdot\sigma(r,X_r^\nu,\nu_r)dW_r. \end{aligned}$$

Proposición 1.2.1. *Supongamos que la función de valor V es de clase $\mathcal{C}^{1,2}([0,T),\mathbb{R}^d)$. Sean $k(\cdot,\cdot,u)$ y $f(\cdot,\cdot,u)$ continuas en (t,x) para $u \in U$ fijo y g semicontinua por abajo. Además, supongamos que k es acotada por abajo. Entonces, para todo $(t,x) \in \boldsymbol{S}$,*

$$-\partial_t V(t,x) - H(t,x,V(t,x),DV(t,x),D^2V(t,x)) \geq 0,$$

con condición terminal $V(T,\cdot) \geq g(\cdot)$.

Demostración. Sean $(t,x) \in \mathbf{S}$ y $u \in U$ fijos.
Consideremos el proceso de control constante $\nu = u$ y al proceso controlado X con condicón inical $X_t = x$.
Definimos, para cada $h > 0$, al tiempo de paro θ_h como,

$$\theta_h \coloneqq \text{ínf}\{s > t : (s-t, X_s - x) \notin [0,h) \times \alpha B\},$$

con $\alpha > 0$ constante y B la bola de radio uno con centro en cero.

Notemos que $\theta_h \to t$, $\mathbb{P}$ casi seguramente cuando $h \downarrow 0$. Y además $\theta_h = t + h < T$, para $h \leq \hat{h}(\omega)$ suficientemente pequeña.

Ahora, ya que V es continua, por la segunda desigualdad del Teorema 1.2.2 se sigue que

$$V(t,x) \geq \sup_{v\in\mathcal{U}_t} \mathbb{E}\bigg(\int_t^{\theta_h^v} \beta^v(t,r)f(r,X_r^{t,x,v},v)dr + \beta(t,\theta_v^h)V(\theta_v^h, X_{\theta_h^v}^{t,x,v})\bigg)$$

$$\geq \mathbb{E}\bigg(\int_t^{\theta_h^\nu} \beta^\nu(t,r) f(r, X_r^{t,x,\nu}, \nu)dr + \beta(t,\theta_\nu^h) V(\theta_\nu^h, X_{\theta_h^\nu}^{t,x,\nu})\bigg),$$

entonces,

$$0 \leq \mathbb{E}\bigg(V(t,x) - \int_t^{\theta_h^\nu} \beta^\nu(t,r) f(r, X_r^{t,x,\nu}, \nu)dr - \beta(t,\theta_\nu^h) V(\theta_\nu^h, X_{\theta_h^\nu}^{t,x,\nu})\bigg).$$

Y como $\beta^\nu(t,s) = \exp\big(-\int_t^s k(r, X_r^{t,x,\nu}, \nu_r)dr\big)$, se tiene que

$$0 \leq \mathbb{E}\bigg(\beta^\nu(0,t)V(t,x) - \int_t^{\theta_h^\nu} \beta^\nu(0,r) f(r, X_r^{t,x,\nu}, \nu)dr - \beta(0,\theta_\nu^h) V(\theta_\nu^h, X_{\theta_h^\nu}^{t,x,\nu})\bigg).$$

Por otro lado, por la Observación 1.2.2 se satisface que

$$\begin{aligned}\beta^\nu(0,\theta_\nu^h)V(\theta_\nu^h, X_{\theta_h^\nu}^\nu) - \beta^\nu(0,t)V(t,x) &= \int_t^{\theta_h^\nu} \beta^\nu(0,r)\big(\partial_t V(r, X_r^{t,x,\nu}) + \mathcal{L}^\nu V(r, X_r^{t,x,\nu})\big)dr \\ &\quad + \int_t^{\theta_h^\nu} \beta^\nu(0,r) DV(r, X_r^{t,x,\nu}) \cdot \sigma(r, X_r^{t,x,\nu}) dW_r,\end{aligned}$$

así,

$$\begin{aligned}0 \leq &-\mathbb{E}\bigg(\int_t^{\theta_h^\nu} \beta^\nu(0,r)\big(\partial_t V(r, X_r^{t,x,\nu}) + \mathcal{L}^\nu V(r, X_r^{t,x,\nu}) + f(r, X_r^{t,x,\nu}, \nu)\big)dr\bigg) \\ &-\mathbb{E}\bigg(\int_t^{\theta_h^\nu} \beta^\nu(0,r) DV(r, X_r^{t,x,\nu}) \cdot \sigma(r, X_r^{t,x,\nu}) dW_r\bigg).\end{aligned}$$

Ahora observemos que, al ser $\beta^\nu(0,r)DV(r, X_r^{t,x,\nu}) \cdot \sigma(r, X_r^{t,x,\nu})$ continua, es acotada en el intervalo $[t, \theta_h]$ entonces, la integral estocástica tiene esperanza cero.

Por otra parte, en $[t, \theta_t^\nu]$, la función $\beta^\nu(0,r)\big(\partial_t V(r, X_r^{t,x,\nu}) + \mathcal{L}^\nu V(r, X_r^{t,x,\nu}) + f(r, X_r^{t,x,\nu}, \nu)\big)$ es acotada. Entonces, para alguna constante C, se satisface que

$$-\mathbb{E}\bigg(\frac{1}{h}\int_t^{\theta_h^\nu} \beta^\nu(0,r)\big(\partial_t V(r, X_r^{t,x,\nu}) + \mathcal{L}^\nu V(r, X_r^{t,x,\nu}) + f(r, X_r^{t,x,\nu}\nu)\big)dr\bigg) \leq \frac{C}{h}\mathbb{E}(\theta_h^\nu - t) = C,$$

donde se utilizó que $\theta_h = t + h$ para $h \leq \hat{h}(\omega)$ suficientemente pequeña. Por lo tanto, por convergencia dominada,

$$\begin{aligned}0 &\leq -\lim_{h\downarrow 0} \mathbb{E}\bigg(\frac{1}{h}\int_t^{\theta_h^\nu} \beta^\nu(0,r)\big(\partial_t V(r, X_r^{t,x,\nu}) + \mathcal{L}^\nu V(r, X_r^{t,x,\nu}) + f(r, X_r^{t,x,\nu}, \nu)\big)dr\bigg) \\ &= -\mathbb{E}\bigg(\lim_{h\downarrow 0} \frac{1}{h}\int_t^{\theta_h^\nu} \beta^\nu(0,r)\big(\partial_t V(r, X_r^{t,x,\nu}) + \mathcal{L}^\nu V(r, X_r^{t,x,\nu}) + f(r, X_r^{t,x,\nu}, \nu)\big)dr\bigg).\end{aligned}$$

Finalmente, por el teorema fundamental del cálculo,

$$-\partial_t V(t,x) - \mathcal{L}^{\nu} V(t,x) - f(t,x,\nu) \geq 0,$$

pues $\theta_h^{\nu} \to t$, $\mathbb{P}$-c.s. Como ν fue arbitrario se concluye que

$$-\partial_t V(t,x) - H(t,x,V(t,x),DV(t,x),D^2V(t,x)) \geq 0.$$

$\blacksquare$

Proposición 1.2.2. *Supongamos que V es de clase $\mathcal{C}^{1,2}([0,T),\mathbb{R}^d)$ y $H(\cdot,\cdot,V,DV,D^2V) < \infty$. Supongamos además que k es acotada y la función H es semicontinua por arriba. Entonces, para todo $(t,x) \in \boldsymbol{S}$,*

$$-\partial_t V(t,x) - H(t,x,V(t,x),DV(t,x),D^2V(t,x)) \leq 0,$$

con condición terminal $V(T,\cdot) \leq g(\cdot)$.

Demostración. Procederemos por contradicción. Supongamos que existe $(t_0,x_0) \in \mathbf{S}$ tal que

$$\partial_t V(t_0,x_0) + H(t_0,x_0,V(t_0,x_0),DV(t_0,x_0),D^2V(t_0,x_0)) < 0. \tag{1.2}$$

Para $\epsilon > 0$, definimos a $\phi \geq V$ como

$$\phi(t,x) = V(t,x) + \epsilon(|t-t_0|^2 + |x-x_0|^4).$$

Notemos que

$$\begin{aligned} (V-\phi)(t_0,x_0) = 0, &\quad D(V-\phi)(t_0,x_0) = 0, \\ D^2(V-\phi)(t_0,x_0) = 0, &\quad \partial_t(V-\phi)(t_0,x_0) = 0. \end{aligned}$$

Entonces, al ser H semicontinua por arriba y por (1.2) se satisface que

$$h(t,x) \coloneqq \partial_t \phi(t,x) + H(t,x,\phi,D\phi,D^2\phi) < 0,$$

para $(t,x) \in \mathcal{N}_r$, con $r > 0$ suficientemente pequeño y $\mathcal{N}_r \coloneqq (t_0 - r, t_0 + r) \times rB_{x_0}$. Donde B_{x_0} es la bola abierta de radio uno con centro en x_0.

Definimos a $-\eta \coloneqq \max\limits_{\partial\mathcal{N}_r}(V-\phi)$. Notemos que, por definición de ϕ, $-\eta < 0$. Para $\nu \in \mathcal{U}_{t_0}$, consideremos al tiempo de paro

$$\theta^{\nu} \coloneqq \inf\{t > t_0 : (t, X_t^{t_0,x_0,\nu}) \notin \mathcal{N}_r\},$$

observemos que, por continuidad de las trayectorias de $X_t^{t_0,x_0,\nu}$, $(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu}) \in \partial\mathcal{N}_r$, entonces, $\phi(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu}) > \eta + V(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu})$, por definición de $-\eta$.

Ahora, por la Observación 1.2.2 tenemos que

$$\begin{aligned}\beta^\nu(t_0,t_0)V(t_0,x_0) &= \beta^\nu(t_0,t_0)\phi(t_0,x_0)\\ &= \mathbb{E}(\beta^\nu(t_0,\theta^\nu)\phi(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu}))\\ &\quad - \mathbb{E}\left(\int_{t_0}^{\theta^\nu} \beta^\nu(t_0,r)\big(\partial_t\phi(r,X_r^{t_0,x_0,\nu}) + \mathcal{L}^\nu\phi(r,X_r^{t_0,x_0,\nu})\big)dr\right),\end{aligned}$$

donde utilizamos que $(V-\phi)(t_0,x_0) = 0$. Además, como $\beta^\nu(t_0,t_0) = 1$ y por definición de h se satisface que

$$-h(t,x) + f(t,x,\nu) \le -\big(\partial_t\phi(t,x) + \mathcal{L}^\nu\phi(t,x)\big),$$

entonces

$$\begin{aligned}V(t_0,x_0) \ge&\mathbb{E}(\beta^\nu(t_0,\theta^\nu)\phi(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu}))\\ &+ \mathbb{E}\left(\int_{t_0}^{\theta^\nu} \beta^\nu(t_0,r)\big(f(r,X_r^{t_0,x_0,\nu},\nu_r) - h(r,X_r^{t_0,x_0,\nu})\big)dr\right).\end{aligned}$$

Por otro lado, como $(r, X_r^{t_0,x_0,\nu}) \in \mathcal{N}_r$, en $[t_0,\theta^\nu)$, entonces $h(r,X_r^{t_0,x_0,\nu}) < 0$. Y recordemos que

$$\phi(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu}) > \eta + V(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu}),$$

por lo tanto

$$V(t_0,x_0) \ge \eta\mathbb{E}(\beta^\nu(t_0,\theta^\nu)) + \mathbb{E}\left(\int_{t_0}^{\theta^\nu} \beta^\nu(t_0,r)f(r,X_r^{t_0,x_0,\nu},\nu_r)dr + \beta^\nu(t_0,\theta^\nu)V(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu})\right).$$

Luego, como k es acotada, observemos que

$$\beta^\nu(t_0,\theta^\nu) = \exp\left(-\int_{t_0}^{\theta^\nu} k(s,X_s^{t_0,x_0,\nu})ds\right) \le \exp(-r\|k^+\|_\infty),$$

donde se utilizó la definición de θ^ν. De lo anterior tenemos que

$$V(t_0,x_0) \ge \eta\exp(-r\|k^+\|_\infty) + \mathbb{E}\left(\int_{t_0}^{\theta^\nu} \beta^\nu(t_0,r)f(r,X_r^{t_0,x_0,\nu},\nu_r)dr + \beta^\nu(t_0,\theta^\nu)V(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu})\right),$$

y al ser $\eta > 0$,

$$V(t_0,x_0) > \mathbb{E}\left(\int_{t_0}^{\theta^\nu} \beta^\nu(t_0,r)f(r,X_r^{t_0,x_0,\nu},\nu_r)dr + \beta^\nu(t_0,\theta^\nu)V(\theta^\nu, X_{\theta^\nu}^{t_0,x_0,\nu})\right).$$

Como tomamos $\nu \in \mathcal{U}_{t_0}$ arbitrario, se satisface que

$$V(t_0, x_0) > \sup_{v \in \mathcal{U}_{t_0}} \mathbb{E}\bigg(\int_{t_0}^{\theta^v} \beta^\nu(t_0, r) f(r, X_r^{t_0,x_0,v}, \nu_r) dr + \beta^\nu(t_0, \theta^v) V(\theta^\nu, X_{\theta^v}^{t_0,x_0,v}) \bigg),$$

y como V es continua, lo anterior contradice la Observación 1.2.1 del Teorema 1.2.2. ■

De las Proposiciones 1.2.1 y 1.2.2, podemos concluir el siguiente resultado:

Teorema 1.2.3. *Supongamos que las condiciones de las Proposiciones 1.2.1 y 1.2.2 se satisfacen. Entonces la función de valor V cumple la ecuación de Hamilton-Jacobi-Bellman, esto es,*

$$-\partial_t V(t,x) - H(\cdot, \cdot, V(t,x), DV(t,x), D^2 V(t,x)) = 0, \quad \text{en } \boldsymbol{S},$$

con condición terminal $V(T, \cdot) = g(\cdot)$.

1.2.4. Verificación

Lo crucial en programación dinámica consiste en demostrar que, dada una solución suave a la ecuación de Hamilton-Jacobi-Bellman, este candidato coincide con la función de valor. Tal resultado es llamado *argumento de verificación* y el cual se enuncia a continuación:

Teorema 1.2.4. *Sea $T < \infty$ y $v \in \mathcal{C}^{1,2}([0,T), \mathbb{R}^d) \cap \mathcal{C}([0,T), \mathbb{R}^d)$. Supongamos que $\|k^-\|_\infty < \infty$ y que v y f satisfacen*

$$|f(t,x,u)| + |v(t,x)| \leq C(1 + \|x\|^2 + \|u\|),$$

para alguna constante C y $(t,x,u) \in [0,T) \times \mathbb{R}^d \times U$.

i) Supongamos que para todo $(t,x) \in \boldsymbol{S}$, se satisface

$$-\partial_t v(t,x) - H(t, x, v(t,x), Dv(t,x), D^2 v(t,x)) \geq 0,$$

con condición terminal $v(T, \cdot) \geq g(\cdot)$. Entonces $v \geq V$ en $\boldsymbol{S}$.

ii) Sea v solución de la ecuación HJB, esto es,

$$-\partial_t v(t,x) - H(\cdot, \cdot, v(t,x), Dv(t,x), D^2 v(t,x)) = 0, \quad \text{en } \boldsymbol{S},$$

con condición terminal $v(T, \cdot) = g(\cdot)$. Supongamos además que existe $\hat{u}(t,x)$ minimizador de
$u \mapsto \mathcal{L}^u v(t,x) + f(t,x,u)$, tal que

- *Se satisface*

$$\partial_t v(t,x) + \mathcal{L}^{\hat{u}} v(t,x) + f(t,x,\hat{u}(t,x)) = 0.$$

- *La ecuación diferencial estocástica*

$$dX_s = b(s, X_s, \hat{u}(s, X_s))ds + \sigma(s, X_s, \hat{u}(s, X_s))dW_s,$$

tiene una única solución dada la condición inicial $X_t = x$.

- *El proceso* $\hat{v}_s := \hat{u}(s, X_s)$ *está bien definido en* $\mathcal{U}_0$.

Entonces $v = V$ *y* $\hat{v}$ *es un control óptimo Markoviano.*

Demostración. Demostraremos primero el inciso i). Consideremos $\nu \in \mathcal{U}_0$ un proceso de control arbitrario y X el proceso de estados asociado con $X_t = x$. Definimos, para cada n, al tiempo de paro θ_n como

$$\theta_n := T \wedge \text{ínf}\{s > t : |X_s - x| \geq n\}. \tag{1.3}$$

Ahora, por la Observación 1.2.2 se tiene que

$$\begin{aligned}\beta(t,\theta_n)v(\theta_n, X^{t,x,\nu}_{\theta_n}) - v(t,x) &= \int_t^{\theta_n} \beta(t,r)\big(\partial_t v(r, X^{t,x,\nu}_r) + \mathcal{L}^\nu v(r, X^{t,x,\nu}_r)\big)dr \\ &\quad + \int_t^{\theta_n} \beta(r, X^{t,x,\nu}_r) Dv(r, X^{t,x,\nu}_r) \cdot \sigma(r, X^{t,x,\nu}_r, \nu_r)dW_r. \end{aligned} \tag{1.4}$$

Notemos que $\partial_t v(r, X^{t,x,\nu}_r) + \mathcal{L}^\nu v(r, X^{t,x,\nu}_r) + f(r, X^{t,x,\nu}_r) \leq \partial_t v(r, X^{t,x,\nu}_r) + H(r, X^{t,x,v}_r, v, Dv, D^2 v) \leq 0$, por hipótesis y definición de H. Además, el integrando $\partial_t v(r, X^{t,x,\nu}_r) + \mathcal{L}^\nu v(r, X^{t,x,\nu}_r) \cdot \sigma(r, X^{t,x,\nu}_r, \nu_r)$ es acotado en el intervalo $[t, \theta_n]$, ya que σ y Dv son continuas y $\|k^-\|_\infty < \infty$. Entonces, tomando esperanza a la ecuación anterior,

$$v(t,x) \geq \mathbb{E}\bigg(\int_t^{\theta_n} \beta(t,r) f(r, X^{t,x,\nu}_r, \nu_r)dr + \beta(t,\theta_n) v(\theta_n, X^{t,x,\nu}_{\theta_n})\bigg).$$

Observemos que, como $\|k^-\|_\infty < \infty$, se satisface que

$$\begin{aligned} &\left|\beta(t,\theta_n)v(\theta_n, X^{t,x,\nu}_{\theta_n}) + \int_t^{\theta_n} \beta(t,r) f(r, X^{t,x,\nu}_r, \nu_r)dr\right| \\ &\leq Ce^{T\|k^-\|_\infty}\bigg(1 + \|X^{t,x,\nu}_{\theta_n}\|^2 + \int_t^{\theta_n}(1 + \|\nu_r\|)dr\bigg) \\ &\leq Ce^{T\|k^-\|_\infty}\bigg(1 + \sup_{t\leq s\leq T}\|X^{t,x,\nu}_s\|^2 + \int_t^{\theta_n}(1 + \|\nu_r\|)dr\bigg), \end{aligned}$$

donde se utilizó la hipótesis sobre las funciones f y v. Además, por el Teorema 1.2.1, la expresión anterior es integrable. Entonces, al hacer tender $n \to \infty$, por el teorema de convergencia dominada, de la desigualdad se sigue que

$$\begin{aligned} v(t,x) &\geq \mathbb{E}\Big(\beta(t,T)v(T,X_T^{t,x,\nu}) + \int_t^T \beta(t,x)f(r,X_r^{t,x,\nu},\nu_r)dr\Big) \\ &\geq \mathbb{E}\Big(\beta(t,T)g(X_T^{t,x,\nu}) + \int_t^T \beta(t,x)f(r,X_r^{t,x,\nu},\nu_r)dr\Big) \\ &= J(t,x,\nu), \end{aligned}$$

ya que $v(T,\cdot) \geq g(\cdot)$. Al ser el control $\nu \in \mathcal{U}_0$ arbitrario, tomando supremo sobre los controles $\nu \in \mathcal{U}_0$, de lo anterior se concluye que $v(t,x) \geq V(t,x)$.

Para demostrar *ii*) procederemos de forma análoga. Consideremos $\nu \in \mathcal{U}_0$ arbitrario y X el proceso de estados asociado con $X_t = x$. Tomando θ_n como en (1.3), análogamente al caso anterior, por la fórmula de Itô llegamos a que $v(t,x) \geq V(t,x)$.

Ahora, si tomamos $\hat{\nu}_s = \hat{u}(s,X_s)$, de (1.4) se tiene que

$$\begin{aligned} \beta(t,\theta_n)v(\theta_n,X_{\theta_n}^{t,x,\hat{\nu}}) - v(t,x) = &\int_t^{\theta_n} \beta(r,X_r^{t,x,\hat{\nu}})Dv(r,X_r^{t,x,\hat{\nu}}) \cdot \sigma(r,X_r^{t,x,\hat{\nu}},\hat{\nu}_r)dW_r \\ &- \int_t^{\theta_n} \beta(t,r)f(r,X_r^{t,x,\hat{\nu}},\hat{\nu}_r)dr, \end{aligned}$$

donde se utilizó la hipótesis $\partial_t v(t,x) + \mathcal{L}^{\hat{\nu}}v(t,x) + f(t,x,\hat{\nu}_t) = 0$.

Así, por el mismo argumento anterior, tomando esperanza y por convergencia dominada, se sigue que $v(t,x) = J(t,x,\hat{\nu})$. Y por definición de $V(\cdot)$, $J(t,x,\hat{\nu}) \leq V(t,x)$, con lo que se concluye la prueba. ■

Los resultados anteriores sugieren la siguiente técnica para resolver un problema de control estocástico:

Consideremos la ecuación de Hamilton-Jacobi-Bellman, dada por,

$$-\partial_t v(t,x) - \sup_{u\in U}\{\mathcal{L}^u v(t,x) + f(t,x,u(t,x))\} = 0,$$

para todo $(t,x) \in \mathbf{S}$ y con condición terminal $v(T,x) = g(x)$.

1. Para $(t,x) \in \mathbf{S}$ fijos, resolver $\sup_{u\in U}\{\mathcal{L}^u v(t,x) + f(t,x,u(t,x))\}$ como un problema de maximización sobre $u \in U$. Denotar por $\hat{u}(t,x)$ a tal máximo.

2. Sustituir $\hat{u}(t,x)$ en la ecuación HJB, lo cual resultará en una ecuación diferencial parcial no lineal.

3. Si se encuentra una solución suave a tal ecuación diferencial parcial no lineal, digamos v, entonces v es la solución al problema de control y $\hat{u}(t,x)$ el control óptimo Markoviano.

Tal técnica será la empleada en la demostración de los resultados expuestos en los Capítulos 3 y 4.

1.3. Optimización de portafolios

El problema de optimización de portafolios consiste en hallar la estrategia óptima que maximice la utilidad esperada de un inversionista. Este ha sido un problema clásico en control estocástico y fuente de investigación desde que fue propuesto por Merton (1971) en su trabajo [10].

Problema de optimización de portafolios

Antes de presentar el problema de optimización introduciremos de manera somera dos conceptos importantes dentro de la teoría de portafolios, las *estrategias autofinanciadas* y las *funciones de utilidad*.

- **Estrategias autofinanciadas**

Las estrategias autofinanciadas están estrechamente relacionadas con el concepto de *arbitraje*, que se refiere a la posibilidad de generar una ganancia segura sin contraer riesgo financiero. En este sentido, un mercado con ausencia de arbitraje busca replicar el valor de un producto derivado mediante la negociación simultánea en uno o más mercados, típicamente, en un activo con riesgo y en una cuenta de mercado de dinero, o también conocida como activo sin riesgo [14].

Para poder definir a una estrategia autofinanciada, consideremos un portafolio que consta de una inversión en la cuenta de mercado de dinero a tasa constante $r \geq 0$ y π_t acciones de un activo con riesgo, S_t. Si denotamos por X_t al valor de tal portafolio al tiempo t, notemos que la posición del inversionista en la cuenta de mercado de dinero está dada por $X_t - \pi_t S_t$. Así, la ganancia del inversionista al tiempo t es

$$dX_t = \pi_t dS_t + r(X_t - \pi_t S_t)dt. \tag{1.5}$$

Por otro lado, si denotamos por $M_t = e^{rt}$ al valor de una inversión en la cuenta de mercado de dinero y Γ_t a la posición invertida en dicha cuenta, entonces

$$X_t = \pi_t S_t + \Gamma_t M_t, \tag{1.6}$$

y por la fórmula del producto,

$$dX_t = S_t d\pi_t + \pi_t dS_t + M_t d\Gamma_t + r\Gamma_t M_t dt. \tag{1.7}$$

Luego, si sustituimos (1.6) en (1.5)

$$dX_t = \pi_t dS_t + r\Gamma_t M_t dt, \tag{1.8}$$

por lo tanto, de (1.7) y (1.8) se sigue que

$$S_t d\pi_t + M_t d\Gamma_t = 0. \tag{1.9}$$

Al estar bajo un mercado con ausencia de arbitraje, a cada tiempo t el valor del portafolio debe ser igual al valor del derivado en cuestión, lo cual nos lleva a tener que rebalancear el portafolio a cada tiempo t. De forma precisa, en la ecuación (1.9), el primer término indica el costo por rebalancear la posición en el activo con riesgo y el segundo el costo asociado al rebalancear la posición en la cuenta de mercado de dinero. Notemos que si la suma de estos dos términos no es cero, entonces habría oportunidad de arbitraje; pues si ésta fuera positiva, el dinero puede retirarse como una ganancia y si fuera negativa habría que pagar para poder entrar en la posición del portafolio deseado. Por lo tanto, diremos que una estrategia es ***autofinanciada*** si cumple la ecuación (1.9).

- **Funciones de utilidad**

Las preferencias que tiene un agente al momento de hacer una inversión es un elemento importante a considerar al momento de analizar un portafolio. Tal toma de decisiones está relacionada con la aversión al riesgo de cada inversionista. Esto es, qué tan conservador o arriesgado es un agente ante el riesgo financiero [7]. Este aspecto se ve reflejado en los modelos de inversión mediante las *funciones de utilidad*, que se definen a continuación.

Definición 1.3.1. *Diremos que $U : \mathbb{R}^+ \to \mathbb{R}$, con $U \in \mathcal{C}^1$ es una función de utilidad si es estrictamente concava, estrictamente creciente y satisface que*

$$U'(0) := \lim_{x \downarrow 0} U'(x) = +\infty, \quad U'(\infty) := \lim_{x \to \infty} U'(x) = 0.$$

La función *potencia* es una de las funciones de utilidad más comunes y está definida por

$$U(x) = \begin{cases} \frac{x^\gamma - 1}{\gamma}, & x > 0, \\ -\infty, & x \leq 0, \end{cases}$$

con $\gamma \in (0, 1)$, al cual se le conoce como coeficiente de aversión al riesgo. El caso $\gamma = 0$ corresponde a la función de utilidad *logarítmica* y está dada por

$$U(x) = \begin{cases} \log(x), & x > 0, \\ -\infty, & x \leq 0. \end{cases}$$

Para el desarrollo de este trabajo se utilizará como función de utilidad esta última.

- **Problema de optimización**

Una vez descritos los conceptos anteriores, podemos definir formalmente el problema de optimización. Sea $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ un espacio de probabilidad filtrado, donde $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ satisface las condiciones usuales y consideremos un mercado financiero con un activo sin riesgo, que denotaremos por S^0 y un activo con riesgo, S, donde

$$\begin{aligned} dS_t^0 &= S_t^0 r dt, \\ dS_t &= S_t(\alpha d_t + \sigma_m dW_t), \end{aligned}$$

con $r \geq 0, \alpha \in \mathbb{R}, \sigma_m > 0$ constantes y W un movimiento browniano unidimensional definido en $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Representaremos por $\pi = \{\pi_t, t \in [0, T]\}$ a la cantidad de dinero invertida por el inversionista en el activo S y diremos que el proceso π es una estrategia admisible si es autofinanciada, progresivamente medible respecto a $\mathbb{F}$ y cuadrado integrable. Así, el valor del portafolio asociado a tal estrategia estará dado por

$$\begin{aligned} dX_t^\pi &= \pi_t \frac{dS_t}{S_t} + (X_t^\pi - \pi_t)\frac{dS_t^0}{S_t^0} \\ &= (rX_t + (\alpha - r)\pi_t)dt + \sigma_m \pi_t dW_t. \end{aligned} \tag{1.10}$$

Definimos al problema de optimización de portafolios como sigue

$$V(t, x) := \sup_{\pi \in \mathcal{U}_0} \mathbb{E}(U(X_T^{t,x,\pi})), \tag{1.11}$$

donde $X^{t,x,\pi}$ es la solución a (1.10) con condición inicial $X_t^{t,x,\pi} = x$, $U(\cdot)$ es la función de utilidad asociada al inversionista y $\mathcal{U}_0$ el conjunto de estrategias admisibles. Como ya se mencionó, el objetivo es hallar la estrategia óptima π^* que maximiza la utilidad esperada de un inversionista.

Notemos que, utilizando el Teorema 1.2.3, la ecuación de Hamilton-Jacobi-Bellaman se ve de la siguiente forma

$$-\partial_t V(t, x) - \sup_{\pi \in \mathcal{U}_0} \mathcal{L}^\pi V(t, x) = 0,$$

donde el operador $\mathcal{L}^{\pi}V(t,x)$ está dado por

$$\mathcal{L}^{\pi}V(t,x) = \frac{\partial V(t,x)}{\partial x}\big(rx + (\alpha - r)\pi\big) + \frac{1}{2}\frac{\partial^2 V(t,x)}{\partial x^2}\sigma_m^2\pi^2,$$

pues para este caso en particular las funciones $f(t,x,\pi)$ y $k(t,x,\pi)$ del Teorema 1.2.3 son cero.

El enfoque bajo el cual se plantea el problema de optimización de portafolios se ha ido adaptando a las distintas preferencias de los inversionistas. Para objetivos de este trabajo será de relevancia considerar el modelo CAPM y las estrategias tipo convergentes, conceptos que se abordan en el siguiente capítulo.

Capítulo 2

El modelo

En este capítulo expondremos el trabajo desarrollado en [2], donde estudian el problema de optimización, y se presentará el contexto financiero de donde emerge tal modelo, como son el modelo CAPM y las estrategias convergentes. El modelo planteado por Sühan et al. en su trabajo [2] es una modificación al planteado por Liu y Timmermann en [8], la cual radica en suponer que los *errores* en los precios asociados a los activos están regidos por una cadena de Markov a tiempo continuo y con espacio de estados discreto, también conocido como *cambio de régimen.*

2.1. Modelo CAPM

El modelo CAPM, por sus siglas en inglés Capital Asset Pricing Model, busca describir la relación entre el riesgo sistemático y el rendimiento esperado de un activo. De forma más precisa, si denotamos por r a la tasa libre de riesgo de mercado y α_m al rendimiento esperado del mercado, entonces el riesgo esperado de un activo está dado por

$$\alpha = r + \beta(\alpha_m - r) = r + \beta\mu_m,$$

donde μ_m denota la prima de riesgo de mercado y β la beta del activo, términos que se explican a continuación: si un inversionista tiene una posición en el índice de mercado, además de invertir en un activo sin riesgo a tasa r, obtendrá un rendimiento esperado extra, $\alpha_m - r$, a éste se le conoce como *prima de riesgo de mercado.* Por otro lado, el parámetro β describe la volatilidad de un activo en comparación con el comportamiento del mercado. Por ejemplo, si $\beta = 1$ querría decir que el activo en cuestión se mueve en promedio igual que el mercado. Los activos con $\beta > 1$ se consideran como *agresivos*, ya que implican mayor riesgo en la inversión; mientras que si $\beta < 1$, entonces se dice que son activos *defensivos* [11].

Es de interés para los inversionistas el contar con una estrategia que, independientemente de si el mercado suba o baje, sea neutral al riesgo sistemático. Diremos que una estrategia es *neutral al riesgo de mercado* o *beta neutral*, si la beta asociada es cero, es decir, el portafolio no tiene correlación ante movimientos del mercado y así el rendimiento esperado será igual a la tasa libre de riesgo [11].

Problema de optimización bajo el modelo CAPM

Bajo el modelo CAPM la dinámica del activo con riesgo está dada por

$$dS_t = S_t\big((r + \beta\mu_m)dt + \sigma_m dB_t\big),$$

por lo que el valor del portafolio W asociado a la estrategia π es

$$dW_t^\pi = \big(rX_t + \beta\mu_m\pi_t\big)dt + \sigma_m\pi_t dB_t.$$

Mientras que el problema de control está definido como en (1.11). Además, en la ecuación HJB el operador $\mathcal{L}^\pi V(t,x)$ está dado por

$$\mathcal{L}^\pi V(t,w) = \frac{\partial V(t,w)}{\partial w}\big(rw + \beta\mu_m\pi\big) + \frac{1}{2}\frac{\partial^2 V(t,w)}{\partial w^2}\sigma_m^2\pi^2.$$

Una de las técnicas para resolver el problema de optimización de portafolios es mediante control estocástico, particularmente como se explicó en la Sección 1.2.

Estrategias tipo convergente

Cuando hay divergencias en el precio de un activo respecto al precio justo se dice que existe un *error* en el precio. Las estrategias de tipo convergente buscan explotar *errores* temporales en los precios de dos activos que tendrán un *payoff* igual (o similar) en el futuro [8].

Un ejemplo típico en este ámbito es el *trading a pares*, que consiste en simultáneamente comercializar activos cointegrados mediante la compra de aquellos subestimados y vender en corto los sobrevalorados, con la expectativa de que en alguna fecha futura los errores se habrán ajustado, beneficiándose así de la convergencia de los precios.

La siguiente gráfica muestra el histórico de precios de las acciones de Ford (F) y General Motors (GM)[1]. Esta imagen nos permite describir visualmente el objetivo de una estrategia *a pares*. Como se puede apreciar estas dos acciones muestran correlación en el histórico de sus precios y

[1] Fuente: https://finance.yahoo.com/

además, en el mes de octubre de 2014 -aproximadamente- podemos ver que los activos comienzan a comportarse en direcciones opuestas para volver a un comportamiento similar en abril de 2015. El objetivo es aprovechar el *spread* (diferencia del logaritmo de los precios) generado al presentar *errores* en los precios de estos activos, beneficiándose de la convergencia a futuro de éstos.

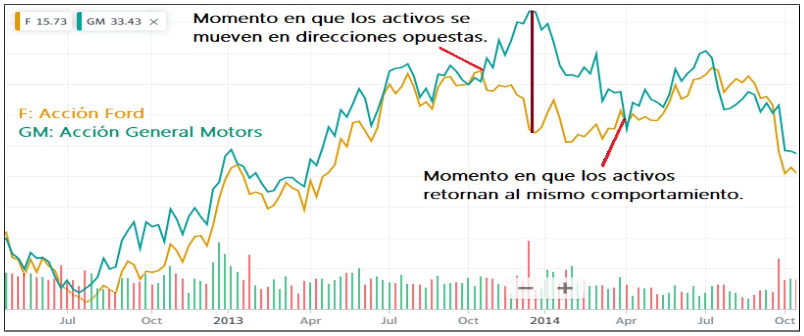

Figura 2.1: Histórico de precios de las acciones de Ford (F) y General Motors (GM).

Para ejemplificar, bajo el modelo CAPM, consideremos el portafolio de inversión descrito en [8], esto es, un mercado con un activo sin riesgo que ofrece una tasa de rendimiento constante $r \geq 0$ y los siguientes activos con riesgo:

- Índice de mercado, el cual seguirá la siguiente dinámica

$$\frac{dS_t^m}{S_t^m} = (r + \mu_m)dt + \sigma_m dB_t^m.$$

- Un par de activos cointegrados con precios dados por

$$\begin{aligned}
\frac{dS_t^1}{S_t^1} &= (r + \beta\mu_m - \lambda_1 X_t)dt + \beta\sigma_m dB_t^m + \sigma dB_t + bdB_t^1, \\
\frac{dS_t^2}{S_t^2} &= (r + \beta\mu_m + \lambda_2 X_t)dt + \beta\sigma_m dB_t^m + \sigma dB_t + bdB_t^2,
\end{aligned}$$

donde $\sigma > 0$ y $b > 0$ son constantes, $\beta \in \mathbb{R}$ representa la *beta* de los activos S_1 y S_2, $\mu_m \in \mathbb{R}$ es la prima de riesgo de mercado, $\sigma_m > 0$ la volatilidad del mercado, $r + \beta\mu_m - \lambda_i X_t$ denotan el rendimiento de los activos S_i, $i = 1, 2$, $\sigma dB_t + bdB_t^i$ el riesgo asociado al sector financiero para cada S_i, $i = 1, 2$, y finalmente, (B, B^1, B^2) es un movimiento browniano tridimensional e independiente del movimiento browniano B^m. Por otro lado, el *spread* de los precios está dado por el proceso $X = \log S^1 - \log S^2$.

Notemos que si, para alguna $i = 1, 2$, $\lambda_i = 0$, entonces el rendimiento esperado del activo correspondiente satisface el modelo CAPM y por ende éste no presenta *errores* en el precio.

Denotaremos por W^π al valor del portafolio asociado a la estrategia admisible $\pi = (\pi^m, \pi^1, \pi^2)$, donde π^m, π^1 y π^2 representan las fracciones de dinero invertido en S^m, S^1 y S^2, respectivamente. Se sigue entonces que la dinámica de W, es decir, el capital ganado debido a cambios en los activos S^m, S^1 y S^2, más los intereses generados, está dada por

$$\begin{aligned}\frac{dW_t^\pi}{W_t^\pi} &= \pi_t^m \frac{dS_t^m}{S_t^m} + \pi_t^1 \frac{dS_t^1}{S_t^1} + \pi_t^2 \frac{dS_t^2}{S_t^2} + r\big(1 - \pi_t^m - \pi_t^1 - \pi_t^2\big)dt \\ &= rdt + (\pi_t^m + \beta(\pi_t^1 + \pi_t^2))(\mu_m dt + \sigma_m dW_t) \\ &\quad + \pi_t^1(\sigma dB_t + bdB_t^1 - \lambda_1 X_t dt) + \pi_t^2(\sigma dB_t + bB_t^1 + \lambda_2 X_t dt).\end{aligned}$$

La ecuación de Hamilton-Jacobi-Bellman se ve de la siguiente forma

$$-\partial_t V(t,w,x) - \sup_{\pi \in \mathcal{U}_0} \mathcal{L}^\pi V(t,w,x) = 0,$$

donde $V(t,w,x) = \sup_{\pi \in \mathcal{U}_0} \mathbb{E}(U(W_T^{t,w,x,\pi}))$ y el operador $\mathcal{L}^\pi V(t,w,x)$ se define como

$$\begin{aligned}\mathcal{L}^\pi V(t,w,x) = &-(\lambda_1 + \lambda_2)xV_x(t,w,x) + b^2 V_{xx}(t,w,x) + b^2(\pi^1 - \pi^2)wV_{xw}(t,x,w) \\ &+ wV_w(t,w,x)\big[r + \mu_m[\pi^m + \beta(\pi^1 + \pi^2)] + x(\lambda_2\pi^2 - \lambda_1\pi^1)\big] \\ &+ \frac{w^2}{2}V_{ww}(t,x,w)\left(\sigma_m^2[\pi^m + \beta(\pi^1 + \pi^2)]^2 + \sigma^2(\pi^1 + \pi^2) + (b\pi^1)^2 + (b\pi^2)^2\right),\end{aligned}$$

con V_x, V_{xx}, V_w, V_{xw} y V_{ww} las derivadas parciales de $V(t,w,x)$.

En el modelo descrito en [2] se utlizan estrategias del tipo convergente, incluyendo además el caso en que son *beta neutrales*. Tal modelo es una modificación al que presentamos anteriormente propuesto por Liu y Timmermann en [8], en la siguiente sección se explica a detalle tal modelo.

2.2. Descripción del modelo

El comportamiento del mercado financiero tiende a cambiar abruptamente, adicionalemente, después de dicho cambio, el nuevo comportamiento de las variables financieras suele persistir durante varios períodos. Tal fenómeno es el que se aborda al emplear un modelo de cambio de régimen. En [8] se estudia la dinámica de asignación óptima de portafolios maximizando la utilidad esperada terminal vía dos activos cointegrados, con *errores* en los precios, y un índice de mercado. La extensión que se hace en [2] consta en emplear un cambio de régimen en los errores de los precios, o *alfas* como se conoce en la literatura financiera, el cual supone que el *alfa* de los activos cointegrados está regida por una cadena de Markov.

Más precisamente, consideremos $(\Omega, \mathcal{G}, \mathbb{P})$ un espacio de probabilidad y $\mathbb{G} = \{\mathcal{G}_t, t \in [0,T]\}$ una filtración que cumple las condiciones usuales, donde T es finito y coincidirá con el tiempo terminal de la inversión. La filtración $\mathbb{G}$ representará todo el flujo de información y los procesos que se definan de aquí en adelante serán adaptados a ésta.

Sea Y una cadena de Markov a tiempo continuo con espacio de estados $\mathcal{E} = \{e_1, \ldots, e_k\}$, para $k \geq 2$. Denotaremos por $Q = (q^{ij})_{i,j\in\mathcal{E}}$ a la matriz infinitesimal de Y. Consideraremos un mercado con las siguientes características:

Se podrá invertir en un activo sin riesgo, el cual tendrá una tasa de retorno constante $r \geq 0$, y tres activos con riesgo:

- $S^{(m)}$ que representará el índice de mercado (*benchmark*) y para el cual supondremos que la dinámica del precio está dada por:

$$dS_t^{(m)} = S_t^{(m)}[(r+\mu_m)dt + \sigma_m dB_t^{(m)}], \tag{2.1}$$

  donde $\mu_m \in \mathbb{R}$ es la prima de riesgo de mercado, $\sigma_m > 0$ la volatilidad de mercado y $B^{(m)}$ un movimiento browniano estándar.

- $S^{(1)}$ y $S^{(2)}$, que denotan los activos cointegrados y sus precios se modelarán como:

$$\frac{dS_t^{(1)}}{S_t^{(1)}} = (r+\beta_1\mu_m)dt - \lambda_1(Y_t)(X_t - \alpha_1(Y_t))dt + \beta_1\sigma_m dB_t^{(m)} + \sigma dB_t^{(0)} + b_1 dB_t^{(1)}, \tag{2.2}$$

$$\frac{dS_t^{(2)}}{S_t^{(2)}} = (r+\beta_2\mu_m)dt + \lambda_2(Y_t)(X_t - \alpha_2(Y_t))dt + \beta_2\sigma_m dB_t^{(m)} + \sigma dB_t^{(0)} + b_2 dB_t^{(2)}, \tag{2.3}$$

  con $\beta_1 \in \mathbb{R}$, $\beta_2 \in \mathbb{R}$, $b_1 > 0$, $b_2 > 0$ y $\sigma > 0$ constantes, donde $(B^{(0)}, B^{(1)}, B^{(2)})$ es un movimiento browniano tridimensional e independiente de $B^{(m)}$.

Supondremos además que la cadena de Markov Y será independiente del vector $(B^{(m)}, B^{(0)}, B^{(1)}, B^{(2)})$ y que las funciones $\lambda(\cdot)$ y $\alpha(\cdot)$ son localmente Lipschitz continuas y con crecimiento lineal.

Los parámetros en (2.2) y (2.3) se describen enseguida:

- el proceso X definido por $X_t = \log S_t^{(1)} - \log S_t^{(2)}$ modelará el *spread* de los precios. Además, por la fórmula de Itô dada en el Teorema 1.1.4, se puede ver que

$$\begin{aligned} dX_t =& (r+\beta_1\mu_m)dt - \frac{1}{2}\big(\beta_1^2\sigma_m^2 + \sigma^2 + b_1^2\big)dt - \lambda_1(Y_t)(X_t - \alpha_1(Y_t))dt + \beta_1\sigma_m dB_t^{(m)} + \sigma dB_t^{(0)} \\ &+ b_1 dB_t^{(1)} - (r+\beta_2\mu_m)dt + \frac{1}{2}\big(\beta_2^2\sigma_m^2 + \sigma^2 + b_2^2\big)dt - \lambda_2(Y_t)(X_t - \alpha_2(Y_t))dt - \beta_2\sigma_m dB_t^{(m)} \\ &- \sigma dB_t^{(0)} - b_2 dB_t^{(2)}, \end{aligned}$$

por lo que

$$\begin{aligned}dX_t =&\big[\Gamma_1 - \lambda_1(Y_t)(X_t - \alpha_1(Y_t)) - \lambda_2(Y_t)(X_t - \alpha_2(Y_t))\big]dt\\&+ (\beta_1 - \beta_2)\sigma_m dB_t^{(m)} + b_1 dB_t^{(1)} - b_2 dB_t^{(2)}, \qquad X_0 \in \mathbb{R},\end{aligned}$$

con $\Gamma_1 = (\beta_1 - \beta_2)\mu_m - \frac{1}{2}((\beta_1^2 - \beta_2^2)\sigma_m^2 + b_1^2 - b_2^2)$.

- $(r + \beta_1\mu_m)dt - \lambda_1(Y_t)(X_t - \alpha_1(Y_t))dt$ y $(r + \beta_2\mu_m)dt - \lambda_2(Y_t)(X_t - \alpha_2(Y_t))dt$ representan el rendimiento esperado *infinitesimal* del activo $S^{(1)}$ y $S^{(2)}$, respectivamente.
- La exposición al riesgo de mercado[2] para los activos $S^{(1)}$ y $S^{(2)}$ estará modelada por $\beta_1\sigma_m dB_t^{(m)}$ y $\beta_2\sigma_m dB_t^{(m)}$, respectivamente.
- El riesgo inherente al sector financiero de cada uno de los activos se representa por $\sigma dB_t^{(0)} + b_1 dB_t^{(1)}$ y $\sigma dB_t^{(0)} + b_2 dB_t^{(2)}$. A este riesgo se le conoce como *diversificable* o riesgo *idiosincrático* y se refiere a aquel tipo de riesgo que afecta exclusivamente a un sector en concreto o a un grupo de acciones.

Algunas observaciones del modelo

1. Liu y Timmermann en su trabajo [8] suponen que los activos en los que se invierte tienen el mismo *payoff* pero se comercializan a distintos precios, notemos que si tomamos $\beta_1 = \beta_2$ y $b_1 = b_2$ rescatamos el modelo descrito en [8] pero con cambio de régimen.
2. Considerar los parámetros λ_1 y λ_2 dependientes de la cadena de Markov Y implica que los errores en los precios dependen de algún factor económico o de mercado que no puede ser directamente observable por el inversionista, esto es, un cambio de régimen en el *alfa* de los activos. Esto último nos permite obtener un modelo más realista que se ajusta al comportamiento visto en *tranding convergente* [2].
3. Si $\lambda_i(Y_t)(X_t - \alpha_i(Y_t)) = 0$, para alguna $i = 1, 2$, entonces los rendimientos en los precios satisfacen el modelo CAPM. Es decir, el rendimiento esperado de los activos dependerá únicamente de la tasa libre de riesgo y de la prima de riesgo de mercado, módulo la correlación entre el activo subyacente y el índice de mercado.
4. Al permitir que las *betas* de los activos sean diferentes es posible plantear el problema de optimización bajo la perspectiva *beta neutral*. Tal enfoque es una segunda extensión al modelo de Liu y Timmermann en [8].

[2]Medida de una futura pérdida potencial resultante de un evento en específico.

Capítulo 3

Optimización bajo información completa

En este capítulo se estudiará el problema de optimización suponiendo que la cadena de Markov Y, que rige el *alfa* de los activos cointegrados, es observable por el inversionista. Se planteará el problema de control a resolver y finalmente se identificará la estrategia óptima. Se caracterizará también la función de valor como la única solución a un sistema de ecuaciones diferenciales ordinarias derivado de la ecuación de Hamilton-Jacobi-Bellman. Se aborda primero el caso en el que no se tiene restricción sobre la asignación de los activos, el cual llamaremos *clásico*. La segunda parte corresponde al caso *beta neutral*, donde se restringe el problema a portafolios que sean neutrales al riesgo de mercado.

3.1. Caso clásico

Definición 3.1.1. *Una estrategia $h = (h^{(m)}, h^{(1)}, h^{(2)})$ es $\mathbb{G}$-admisible si es una estrategia auto-financiada y $\mathbb{G}$-predecible tal que*

$$\mathbb{E}\left(\int_0^T (h_t^{(m)})^2 + (h_t^{(1)})^2 + (h_t^{(2)})^2 dt\right) < \infty. \tag{3.1}$$

Denotaremos por $\mathcal{A}$ al conjunto de estrategias $\mathbb{G}$-admisibles.

En lo que sigue, W^h representará el valor del portafolio asociado a la estrategia $h = (h^{(m)}, h^{(1)}, h^{(2)})$. Las cantidades $h_t^{(m)}, h_t^{(1)}$ y $h_t^{(2)}$ representan las fracciones de dinero invertido al tiempo $t \in [0, T]$ en el índice de mercado $S^{(m)}$ y en los activos cointegrados $S^{(1)}$ y $S^{(2)}$, respectivamente.

La dinámica de W^h estará dada por el capital ganado debido a cambios en los activos $S^{(m)}, S^{(1)}$

y $S^{(2)}$, más los intereses generados, esto es

$$dW_t^h = W_t^h\left[h_t^{(1)}\frac{dS_t^{(1)}}{S_t^{(1)}} + h_t^{(2)}\frac{dS_t^{(2)}}{S_t^{(2)}} + h_t^{(m)}\frac{dS_t^{(m)}}{S_t^{(m)}}\right] + rW_t^h\big[1 - h_t^{(1)} - h_t^{(2)} - h_t^{(m)}\big]dt,$$

sustituyendo (2.1), (2.2) y (2.3), obtenemos

$$\begin{aligned} dW_t^h = W_t^h\Big[&\big(r + (h_t^{(m)} + h_t^{(1)}\beta_1 + h_t^{(2)}\beta_2)\mu_m + h_t^{(2)}\lambda_2(Y_t)(X_t - \alpha_2(Y_t)) \\ &- h_t^{(1)}\lambda_1(Y_t)(X_t - \alpha_1(Y_t))\big)dt + \sigma(h_t^{(1)} + h_t^{(2)})dB_t^{(0)} + b_1h_t^{(1)}dB_t^{(1)} + b_2h_t^{(2)}dB_t^{(2)} \\ &+ \sigma_m(h_t^{(m)} + \beta_1h_t^{(1)} + \beta_2h_t^{(2)})dB_t^{(m)}\Big] \\ = f(t, X_t, Y_t)&W_t^h dt + \sum_{l=0}^{3} g_l(t)W_t^h dB_t^l, \end{aligned}$$

con $f(t, X_t, Y_t) \coloneqq r + (h_t^{(m)} + h_t^{(1)}\beta_1 + h_t^{(2)}\beta_2)\mu_m + h_t^{(2)}\lambda_2(Y_t)(X_t - \alpha_2(Y_t)) - h_t^{(1)}\lambda_1(Y_t)(X_t - \alpha_1(Y_t))$, $(B^0, B^1, B^2, B^3) \coloneqq (B^{(0)}, B^{(1)}, B^{(2)}, B^{(m)})$ y

$$\begin{aligned} g_0(t) &\coloneqq \sigma(h_t^{(1)} + h_t^{(2)}), \\ g_1(t) &\coloneqq b_1h_t^{(1)}, \\ g_2(t) &\coloneqq b_2h_t^{(2)}, \\ g_3(t) &\coloneqq \sigma_m(h_t^{(m)} + \beta_1h_t^{(1)} + \beta_2h_t^{(2)}). \end{aligned}$$

De lo anterior, usando el Lema 1.1.1, para $u \leq t$ se tiene que

$$\begin{aligned} W_t^h &= W_u^h \exp\left(\int_u^t \left[f(s, X_s, Y_s) - \frac{1}{2}\sum_{l=0}^{3} g_l^2(s)\right]ds + \sum_{l=0}^{3}\int_u^t g_l(s)dB_s^l\right) \\ &\eqqcolon W_u^h \exp(G(u, t, X_s, Y_s; u \leq s \leq t)), \end{aligned} \tag{3.2}$$

donde

$$G(u, t, X_s, Y_s; u \leq s \leq t) \coloneqq \int_u^t \left[f(s, X_s, Y_s) - \frac{1}{2}\sum_{l=0}^{3} g_l^2(s)\right]ds + \sum_{l=0}^{3}\int_u^t g_l(s)dB_s^l.$$

3.1.1. Problema de control

Antes de definir el problema de control introduciremos el siguiente operador:

Definición 3.1.2. *Sea $F(t, w, x, i)$ en $C^{1,2,2}([0, T], \mathbb{R}^+, \mathbb{R})$, para cada $i \in \mathcal{E} = \{e_1, \ldots, e_k\}$, defi-*

nimos al operador de segundo orden del proceso (W, X, Y) como

$$
\begin{aligned}
&\mathcal{L}^h_{\mathbb{G}} F(t,w,x,i) \\
&:= F_x(t,w,x,i)\big[\Gamma_1 - \lambda_1^i(x-\alpha_1^i) - \lambda_2^i(x-\alpha_2^i)\big] \\
&\quad + wF_w(t,w,x,i)\big[r + (h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)\mu_m + h^{(2)}\lambda_2^i(x-\alpha_2^i) - h^{(1)}\lambda_1^i(x-\alpha_1^i)\big] \\
&\quad + \frac{1}{2}w^2 F_{ww}(t,w,x,i)\big[\sigma_m^2(h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)^2 + \sigma^2(h^{(1)} + h^{(2)})^2 + (b_1h^{(1)})^2 + (b_2h^{(2)})^2\big] \\
&\quad + wF_{wx}(t,w,x,i)\big[\sigma_m^2(\beta_1-\beta_2)(h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2) + b_1^2h^{(1)} - b_2^2h^{(2)}\big] \\
&\quad + \frac{1}{2}F_{xx}(t,w,x,i)\big[(\beta_1-\beta_2)^2\sigma_m^2 + b_1^2 + b_2^2\big] + \sum_{j\in\mathcal{E}} F(t,w,x,i)q^{ij}.
\end{aligned}
\tag{3.3}
$$

Más adelante se utilizará la fórmula de Itô para relacionar este operador con la ecuación de Hamilton-Jacobi-Bellman asociada al problema de control.

Consideremos un inversionista que desea maximizar su utilidad esperada asignando su riqueza a un mercado como el descrito en el capítulo anterior y con función de utilidad logarítmica.

Formulamos entonces el problema de control de la siguiente manera:

$$\text{maximizar } \mathbb{E}_{t,w,x,i}(\log W_T^h), \text{ sobre todas las estrategias admisibles,}$$

donde $\mathbb{E}_{t,w,x,i}$ denota la esperanza condicional dado que $W_t = w$, $X_t = x$ y $Y_t = e_i$. La función de valor asociada está definida por

$$V(t,w,x,i) := \sup_{h\in\mathcal{A}} \mathbb{E}_{t,w,x,i}(\log W_T^h).$$

Con lo anterior, el resultado principal de este apartado es el siguiente teorema:

Teorema 3.1.1. *Consideremos una inversión con función de utilidad logarítmica. Entonces la estrategia óptima $h_* = (h_*^{(m)}, h_*^{(1)}, h_*^{(2)}) \in \mathcal{A}$ está dada por*

$$
\begin{aligned}
h_*^{(1)}(t,x,i) &= \frac{-\lambda_1^i(x-\alpha_1^i) - \lambda_2^i(x-\alpha_2^i)\varrho_2}{b_1^2 + b_2^2\varrho_2}, \\
h_*^{(2)}(t,x,i) &= \frac{\lambda_2^i(x-\alpha_2^i) + \lambda_1^i(x-\alpha_1^i)\varrho_1}{b_1^2\varrho_1 + b_2^2}, \\
h_*^{(m)}(t,x,i) &= \frac{\mu_m}{\sigma_m^2} - h_*^{(1)}(t,x,i)\beta_1 - h_*^{(2)}(t,x,i)\beta_2,
\end{aligned}
\tag{HO}
$$

donde $\varrho_1 = \frac{\sigma^2}{\sigma^2+b_1^2}$ y $\varrho_2 = \frac{\sigma^2}{\sigma^2+b_2^2}$. Además, la función de valor es de la forma

$$V(t,w,x,i) = \log(w) + m(t,i)x^2 + n(t,i)x + u(t,i),$$

donde las funciones $m(t,i)$, $n(t,i)$ y $u(t,i)$ para cada $i \in \mathcal{E}$ resuelven el siguiente sistema de *ecuaciones diferenciales ordinarias*

$$\begin{aligned} m_t(t,i) - 2(\lambda_1^i + \lambda_2^i)m(t,i) + \sum_{j\in\mathcal{E}} m(t,j)q^{ij} + \Theta_1^i &= 0, \\ n_t(t,i) - (\lambda_1^i + \lambda_2^i)n(t,i) + \sum_{j\in\mathcal{E}} n(t,j)q^{ij} + 2(\Gamma_1 + \lambda_1^i\alpha_1^i + \lambda_2^i\alpha_2^i)m(t,i) - \Theta_2^i &= 0, \\ u_t(t,i) + \sum_{j\in\mathcal{E}} u(t,j)q^{ij} + \Gamma_2 m(t,i) + (\Gamma_1 + \lambda_1^i\alpha_1^i + \lambda_2^i\alpha_2^i)n(t,i) + \Theta_3^i &= 0, \end{aligned}$$

con condiciones terminales $m(T,\cdot) = 0$, $n(T,\cdot) = 0$ y $u(T,\cdot) = 0$, y los parámetros Γ_1, Γ_2, Θ_k^i están dados por

$$\begin{aligned} \Gamma_1 =& (\beta_1 - \beta_2)\mu_m - \frac{1}{2}((\beta_1^2 - \beta_2^2)\sigma_m^2 + b_1^2 - b_2^2), \\ \Gamma_2 =& \sigma_m^2(\beta_1 - \beta_2)^2 + b_1^2 + b_2^2, \\ \Theta_1^i =& \frac{(b_1\lambda_2^i)^2 + (b_2\lambda_1^i)^2 + \sigma^2(\lambda_1^i + \lambda_2^i)^2}{2[(b_1b_2)^2 + \sigma^2(b_1^2 + b_2^2)]}, \\ \Theta_2^i =& \frac{\alpha_1^i\lambda_1^i(\lambda_1^i(b_2^2 + \sigma^2) + \lambda_2^i\sigma^2) + \alpha_2^i\lambda_2^i(\lambda_2^i(b_1^2 + \sigma^2) + \lambda_1^i\sigma^2)}{(b_1b_2)^2 + \sigma^2(b_1^2 + b_2^2)}, \\ \Theta_3^i =& r + \frac{\mu_m^2}{2\sigma_m^2} + \frac{(\alpha_1^i\lambda_1^i b_2)^2 + (\alpha_2^i\lambda_2^i b_1)^2 + \sigma^2(\alpha_1^i\lambda_1^i + \alpha_2^i\lambda_2^i)^2}{2[(b_1b_2)^2 + \sigma^2(b_1^2 + b_2^2)]}. \end{aligned}$$

En lo anterior m_t, n_t y u_t denotan las derivadas parciales respecto a t de las funciones m, n y u, respectivamente.

Demostración.

La idea de la prueba será usar la técnica descrita en la Sección 1.2, esto es, a partir de la ecuación de Hamilton-Jacobi-Bellman obtener la estrategia óptima $h_* = (h_*^{(m)}, h_*^{(1)}, h_*^{(2)})$. Una vez que encontremos $h_* = (h_*^{(m)}, h_*^{(1)}, h_*^{(2)})$ se propone un candidato a la función de valor y finalmente se verifica que, en efecto, ésta coincide con la solución de la ecuación HJB.

1. Estrategia Óptima

Supongamos que la función de valor $V(\cdot,i)$ es de clase $C^{1,2,2}([0,T],\mathbb{R}^+,\mathbb{R})$, esto implica que satisface la ecuación de HJB dada, para cada $i \in \mathcal{E}$, por

$$V_t(t,w,x,i) + \sup_{h\in\mathbb{A}} \mathcal{L}_{\mathbb{G}}^h V(t,w,x,i) = 0, \tag{3.4}$$

con la condición terminal $V(T,\cdot) = \log(w)$, donde $\mathcal{L}_{\mathbb{G}}^h$ denota el generador del proceso (W,X,Y) definido en (3.3). Además, por (3.2) podemos escribir a la función de valor como

$$V(t,w,x,i) := \sup_{h\in\mathcal{A}} \mathbb{E}_{t,w,x,i}\left[\log W_T^h\right]$$

$$
\begin{aligned}
&= \sup_{h\in\mathcal{A}} \mathbb{E}_{t,w,x,i}\big[\log\big(W_t^h \exp\{G(t,T,X_s,Y_s;t\le s\le T)\}\big)\big]\\
&= \sup_{h\in\mathcal{A}} \Big\{\mathbb{E}_{t,w,x,i}[\log W_t^h] + \mathbb{E}_{t,w,x,i}\big[G(t,T,X_s,Y_s;t\le s\le T)\big]\Big\}\\
&= \log(w) + \sup_{h\in\mathcal{A}} \mathbb{E}_{t,w,x,i}\big[G(t,T,X_s,Y_s;t\le s\le T)\big]\\
&=: \log(w) + \nu(t,x,i), \qquad (3.5)
\end{aligned}
$$

para alguna función $\nu(t,x,i)$, con $\nu(T,\cdot,\cdot)=0$.

Sustituyendo (3.3) y (3.5) en (3.4)

$$
\begin{aligned}
\nu_t(t,x,i) + \sup_{h\in\mathcal{A}} \Bigg[&\nu_x(t,x,i)\big[\Gamma_1 - \lambda_1^i(x-\alpha_1^i) - \lambda_2^i(x-\alpha_2^i)\big]\\
&+ r + (h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)\mu_m + h^{(2)}\lambda_2^i(x-\alpha_2^i) - h^{(1)}\lambda_1^i(x-\alpha_1^i)\\
&- \frac{1}{2}\big[\sigma_m^2(h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)^2 + \sigma^2(h^{(1)}+h^{(2)})^2 + (b_1h^{(1)})^2 + (b_2h^{(2)})^2\big]\\
&+ \frac{1}{2}\Gamma_2\nu_{xx}(t,x,i) + \sum_{j\in\mathcal{E}}\nu(t,x,j)q^{ij}\Bigg] = 0, \qquad (3.6)
\end{aligned}
$$

con $\Gamma_2 = (\beta_1-\beta_2)^2\sigma_m^2 + b_1^2 + b_2^2$.

Hallaremos el supremo que aparece en (3.6) obteniendo la estrategia óptima $h_* = (h_*^{(m)}, h_*^{(3)}, h_*^{(2)})$ mediante las condiciones de primer orden. Esto es, derivando el argumento del supremo respecto de $h^{(m)}, h^{(1)}$ y $h^{(2)}$ se obtendrán los puntos críticos. Luego, se comprobará que éstos en efecto son máximos mediante las condiciones de segundo orden.

Por las condiciones de primer, derivando respecto de $h^{(m)}, h^{(1)}$ y $h^{(2)}$, obtenemos, respectivamente

$$
\frac{\mu_m}{\sigma_m^2} - h_*^{(1)}\beta_1 - h_*^{(2)}\beta_2 - {h_*}^{(m)} = 0, \qquad (3.7)
$$

$$
\beta_1\mu_m - \lambda_1^i(x-\alpha_1^i) - \beta_1\sigma_m^2(h_*^{(m)} + h_*^{(1)}\beta_1 + h_*^{(2)}\beta_2) - \sigma^2(h_*^{(1)} + h_*^{(2)}) - h_*^{(1)}b_1^2 = 0, \qquad (3.8)
$$

$$
\beta_2\mu_m + \lambda_2^i(x-\alpha_2^i) - \beta_2\sigma_m^2(h_*^{(m)} + h_*^{(1)}\beta_1 + h_*^{(2)}\beta_2) - \sigma^2(h_*^{(1)} + h_*^{(2)}) - h_*^{(2)}b_2^2 = 0, \qquad (3.9)
$$

De la ecuación (3.7),

$$
h_*^{(m)} = \frac{\mu_m}{\sigma_m^2} - h_*^{(1)}\beta_1 - h_*^{(2)}\beta_2, \qquad (3.10)
$$

sustituyendo (3.10) en (3.8), obtenemos

$$
-h_*^{(1)}(\sigma^2 + b_1^2) = \lambda_1^i(x-\alpha_1^i) + h_*^{(2)}\sigma^2,
$$

de donde se sigue que

$$h_*^{(1)} = \frac{-\lambda_1^i(x-\alpha_1^i) - h_*^{(2)}\sigma^2}{\sigma^2 + b_1^2}. \tag{3.11}$$

Luego, por (3.9)-(3.11),

$$\lambda_2^i(x-\alpha_2^i) + \lambda_1^i(x-\alpha_1^i)\varrho_1 = h_*^{(2)}[\sigma^2(1-\varrho_1) + b_2^2] = h_*^{(2)}(b_1^2\varrho_1 + b_2^2),$$

con $\varrho_1 = \frac{\sigma^2}{\sigma^2+b_1^2}$. Por lo que,

$$h_*^{(2)} = \frac{\lambda_2^i(x-\alpha_2^i) + \lambda_1^i(x-\alpha_1^i)\varrho_1}{b_1^2\varrho_1 + b_2^2}. \tag{3.12}$$

Finalmente, sustituyendo (3.12) en (3.11),

$$\begin{aligned} h_*^{(1)} &= -\frac{\lambda_1^i(x-\alpha_1^i)}{\sigma^2+b_1^2} - \varrho_1\left[\frac{\lambda_2^i(x-\alpha_2^i) + \lambda_1^i(x-\alpha_1^i)\varrho_1}{b_1^2\varrho_1 + b_2^2}\right] \\ &= -\lambda_1^i(x-\alpha_1^i)\left[\frac{b_1^2\varrho_1 + b_2^2 + \varrho_1^2(\sigma^2+b_1^2)}{(\sigma^2+b_1^2)(b_1^2\varrho_1+b_2^2)}\right] - \lambda_2^i(x-\alpha_2^i)\left[\frac{\varrho_1}{b_1^2\varrho_1+b_2^2}\right]. \end{aligned} \tag{3.13}$$

Pero notemos que,

$$\frac{\varrho_1}{b_1^2\varrho_1+b_2^2} = \frac{\sigma^2}{b_1^2\sigma^2 + b_2^2(\sigma^2+b_1^2)} = \frac{\sigma^2}{b_1^2(\sigma^2+b_2^2) + b_2^2\sigma^2} = \frac{\varrho_2}{b_1^2 + b_2^2\varrho_2}, \tag{3.14}$$

con $\varrho_2 = \frac{\sigma^2}{\sigma^2+b_2^2}$. Por otro lado,

$$\begin{aligned} \frac{b_1^2\varrho_1 + b_2^2 + \varrho_1^2(\sigma^2+b_1^2)}{(\sigma^2+b_1^2)(b_1^2\varrho_1+b_2^2)} &= \frac{b_1^2\varrho_1 + b_2^2 + \varrho_1^2(\sigma^2+b_1^2)}{b_1^2(\sigma^2+b_2^2)+b_2^2\sigma^2} \\ &= \frac{b_1^2\varrho_1 + b_2^2 + \varrho_1^2(\sigma^2+b_1^2)}{\sigma^2+b_2^2}\cdot\frac{1}{b_1^2+b_2^2\varrho_2} \\ &= \frac{b_1^2\sigma^2 + b_2^2\sigma^2 + b_2^2b_1^2 + \sigma^4}{(\sigma^2+b_1^2)(\sigma^2+b_2^2)}\cdot\frac{1}{b_1^2+b_2^2\varrho_2} \\ &= \frac{1}{b_1^2+b_2^2\varrho_2}. \end{aligned} \tag{3.15}$$

Entonces, de (3.13), (3.14) y (3.15) se concluye que

$$h_*^{(1)} = \frac{-\lambda_1^i(x-\alpha_1^i) - \lambda_2^i(x-\alpha_2^i)\varrho_2}{b_1^2 + b_2^2\varrho_2}.$$

Comprobaremos ahora las condiciones de segundo orden. Obteniendo las segundas derivadas se sigue que

$$A = (a_{ij})_{\{i,j=1,2,3\}} := \begin{pmatrix} -\sigma_m^2 & -\sigma_m^2\beta_1 & -\sigma_m^2\beta_2 \\ -\beta_1\sigma_m^2 & -\beta_1^2\sigma_m^2 - \sigma^2 - b_1^2 & -\beta_1\beta_2\sigma_m^2 - \sigma^2 \\ -\beta_2\sigma_m^2 & -\beta_1\beta_2\sigma_m^2 - \sigma^2 & -\beta_2^2\sigma_m^2 - \sigma^2 - b_2^2 \end{pmatrix}.$$

Veremos que la matriz de segundas derivadas es definida no positiva, equivalentemente, que se satisfaga

$$\begin{aligned} a_{11} &< 0, \\ \frac{1}{a_{11}}\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} &< 0, \\ \frac{\det(A)}{\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}} &< 0, \end{aligned}$$

donde $\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}$ denota el menor de la matriz A.

Notemos que $\det(A) = -\sigma_m^2[b_1^2b_2^2 + \sigma^2(b_1^2 + b_2^2)] < 0$ y

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} = \sigma_m^2(\sigma_1^2 + b_1^2) > 0,$$

de donde se verifica que $h_* = \left(h_*^{(m)}, h_*^{(1)}, h_*^{(2)}\right)$ en efecto está dada por (HO).

2. Existencia

Para caracterizar la forma de la función de valor sustituiremos la estrategia óptima $h_* = \left(h_*^{(m)}, h_*^{(1)}, h_*^{(2)}\right)$ en la ecuación (3.6).

Observemos primero que, sustituyendo $h_*^{(m)}$ en (3.8) y (3.9), obtenemos lo siguiente

$$\begin{aligned} -\lambda_1^i(x - \alpha_1^i) - \sigma^2(h_*^{(1)} + h_*^{(2)}) - h_*^{(1)}b_1^2 &= 0, \\ \lambda_2^i(x - \alpha_2^i) - \sigma^2(h_*^{(1)} + h_*^{(2)}) - h_*^{(2)}b_2^2 &= 0, \end{aligned}$$

y multiplicando las ecuaciones anteriores por $h_*^{(1)}$ y $h_*^{(2)}$, respectivamente,

$$\begin{aligned} -h_*^{(1)}\lambda_1^i(x - \alpha_1^i) - \sigma^2(h_*^{(1)} + h_*^{(2)})h_*^{(1)} - (h_*^{(1)}b_1)^2 &= 0, \\ h_*^{(2)}\lambda_2^i(x - \alpha_2^i) - \sigma^2(h_*^{(1)} + h_*^{(2)})h_*^{(2)} - (h_*^{(2)}b_2)^2 &= 0, \end{aligned}$$

por lo que

$$\sigma^2(h_*^{(1)}+h_*^{(2)})^2+(h_*^{(1)}b_1)^2+(h_*^{(2)}b_2)^2=h_*^{(2)}\lambda_2^i(x-\alpha_2^i)-h_*^{(1)}\lambda_1^i(x-\alpha_1^i). \tag{3.16}$$

Entonces, sustituyendo $h_*=\big(h_*^{(m)},h_*^{(1)},h_*^{(2)}\big)$ en (3.6),

$$\begin{aligned}&\nu_t(t,x,i)+\nu_x(t,x,i)\big[\Gamma_1-\lambda_1^i(x-\alpha_1^i)-\lambda_2^i(x-\alpha_2^i)\big]+r+\frac{\mu_m^2}{\sigma_m^2}+h_*^{(2)}\lambda_2^i(x-\alpha_2^i)-h_*^{(1)}\lambda_1^i(x-\alpha_1^i)\\ &-\frac{1}{2}\big[\frac{\mu_m^2}{\sigma_m^2}+\sigma^2(h_*^{(1)}+h_*^{(2)})^2+(b_1h_*^{(1)})^2+(b_2h_*^{(2)})^2\big]+\frac{1}{2}\Gamma_2\nu_{xx}(t,x,i)+\sum_{j\in\mathcal{E}}\nu(t,x,j)q^{ij}=0,\end{aligned}$$

y por (3.16),

$$\begin{aligned}&\nu_t(t,x,i)+\nu_x(t,x,i)\big[\Gamma_1-\lambda_1^i(x-\alpha_1^i)-\lambda_2^i(x-\alpha_2^i)\big]+r+\frac{\mu_m^2}{2\sigma_m^2}\\ &+\frac{1}{2}\left[h_*^{(2)}\lambda_2^i(x-\alpha_2^i)-h_*^{(1)}\lambda_1^i(x-\alpha_1^i)\right]+\frac{1}{2}\Gamma_2\nu_{xx}(t,x,i)+\sum_{j\in\mathcal{E}}\nu(t,x,j)q^{ij}=0.\end{aligned}$$

Notemos que,

$$\begin{aligned}h_*^{(2)}\lambda_2^i(x-\alpha_2^i)&=\frac{(\lambda_2^i)^2(\sigma^2+b_1^2)(x^2-2x\alpha_2^i+(\alpha_2^i)^2)+\lambda_1^i\lambda_2^i\sigma^2(x^2-x(\alpha_1^i+\alpha_2^i)+\alpha_1^i\alpha_2^i))}{b_1^2b_2^2+\sigma^2(b_1^2+b_2^2)},\\ -h_*^{(1)}\lambda_1^i(x-\alpha_1^i)&=\frac{(\lambda_1^i)^2(\sigma^2+b_2^2)(x^2-2x\alpha_1^i+(\alpha_1^i)^2)+\lambda_1^i\lambda_2^i\sigma^2(x^2-x(\alpha_1^i+\alpha_2^i)+\alpha_1^i\alpha_2^i))}{b_1^2b_2^2+\sigma^2(b_1^2+b_2^2)}.\end{aligned}$$

Por lo tanto,

$$\begin{aligned}&\nu_t(t,x,i)+\nu_x(t,x,i)\big[\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i-(\lambda_1^i+\lambda_2^i)x\big]+\Theta_1^ix^2-\Theta_2^ix+\Theta_3^i\\ &+\frac{1}{2}\Gamma_2\nu_{xx}(t,x,i)+\sum_{j\in\mathcal{E}}\nu(t,x,j)q^{ij}=0,\end{aligned} \tag{3.17}$$

con Θ_1^i,Θ_2^i y Θ_3^i dadas como en el enunciado del teorema.

Como función de valor proponemos $\nu(t,x,i)=m(t,i)x^2+n(t,i)x+u(t,i)$. Sustituyendo en (3.17) se tiene que, para cada $(t,x)\in[0,T]\times\mathbb{R}$, se cumple

$$\begin{aligned}&m_t(t,i)x^2+n_t(t,i)x+u_t(t,i)+\big[\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i-(\lambda_1^i+\lambda_2^i)x\big][2m(t,i)x+n(t,i)]\\ &\quad+\Theta_1^ix^2-\Theta_2^ix+\Theta_3^i+\Gamma_2m(t,i)+\sum_{j\in\mathcal{E}}q^{ij}[m(t,j)x^2+n(t,j)x+u(t,j)]\\ &=x^2\left[m_t(t,i)-2m(t,i)(\lambda_1^i+\lambda_2^i)+\Theta_1^i+\sum_{j\in\mathcal{E}}q^{ij}m(t,j)\right]\\ &\quad+x\left[n_t(t,i)+2m(t,i)\big[\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i\big]-n(t,i)(\lambda_1^i+\lambda_2^i)-\Theta_2^i+\sum_{j\in\mathcal{E}}q^{ij}n(t,j)\right]\end{aligned}$$

$$+u_t(t,i) + n(t,i)[\Gamma_1 + \lambda_1^i\alpha_1^i + \lambda_2^i\alpha_2^i] + \Theta_3^i + \Gamma_2 m(t,i) + \sum_{j\in\mathcal{E}} q^{ij}u(t,j)$$
$$= 0,$$

esto es; si para cada $i \in \mathcal{E}$, las funciones $m(t,i)$, $n(t,i)$ y $u(t,i)$ resuelven el siguiente sistema de ecuaciones diferenciales ordinarias

$$\begin{aligned}
m_t(t,i) - 2(\lambda_1^i + \lambda_2^i)m(t,i) + \sum_{j\in\mathcal{E}} m(t,j)q^{ij} + \Theta_1^i = 0,\\
n_t(t,i) - (\lambda_1^i + \lambda_2^i)n(t,i) + \sum_{j\in\mathcal{E}} n(t,j)q^{ij} + 2(\Gamma_1 + \lambda_1^i\alpha_1^i + \lambda_2^i\alpha_2^i)m(t,i) - \Theta_2^i = 0,\\
u_t(t,i) + \sum_{j\in\mathcal{E}} u(t,j)q^{ij} + \Gamma_2 m(t,i) + (\Gamma_1 + \lambda_1^i\alpha_1^i + \lambda_2^i\alpha_2^i)n(t,i) + \Theta_3^i = 0,
\end{aligned}$$

con condiciones terminales $m(T,\cdot) = 0, n(T,\cdot) = 0$ y $u(T,\cdot) = 0$; resulta que $\nu(t,x,i) = m(t,i)x^2 + n(t,i)x + u(t,i)$ resuelve la ecuación de HJB. Luego entonces, la función de valor es de la forma $V(t,w,x,i) = \log(w) + \nu(t,x,i) = \log(w) + m(t,i)x^2 + n(t,i)x + u(t,i)$.

3. Verificación.

Para verificar que la función de valor propuesta coincide con la solución de la ecuación HJB utilizaremos el siguiente lema:

Lema 3.1.1. *El proceso definido por*

$$Z_t = \sum_{n\geq 1} \mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{T_n\leq t\}} - \int_0^t \sum_{i\neq j} q^{ij}\mathbb{1}_{\{Y_{r^-}=i\}}dr$$

es $(\mathbb{G},\mathbb{P})$ *martingala.*

Demostración. Por definición de Z_t, es claro que es adaptado a $\mathbb{G}$ y además es integrable.

Ahora, para $s \leq t$ tenemos que

$$\begin{aligned}
&\mathbb{E}\Big(\sum_{n\geq 1} \mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{T_n\leq t\}}\big|\mathcal{G}_s\Big) - \mathbb{E}\Big(\int_0^t \sum_{i\neq j} q^{ij}\mathbb{1}_{\{Y_{r^-}=i\}}dr\big|\mathcal{G}_s\Big) =\\
&\sum_{n\geq 1} \mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{T_n\leq s\}} - \int_0^s \sum_{i\neq j} q^{ij}\mathbb{1}_{\{Y_{r^-}=i\}}dr\\
&\quad + \mathbb{E}\Big(\sum_{n\geq 1} \mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{s<T_n\leq t\}}\big|\mathcal{G}_s\Big) - \mathbb{E}\Big(\int_s^t \sum_{i\neq j} q^{ij}\mathbb{1}_{\{Y_{r^-}=i\}}dr\big|\mathcal{G}_s\Big).
\end{aligned}$$

Pero,

$$\mathbb{E}\Big(\int_s^t \sum_{i\neq j} q^{ij}\mathbb{1}_{\{Y_{r^-}=i\}}dr\big|\mathcal{G}_s\Big) = \int_s^t \sum_{i\neq j}\sum_{l\in\mathcal{E}} \mathbb{1}_{\{Y_s=l\}} q^{ij}\mathbb{P}(Y_{r^-}=i|Y_s=l)dr \tag{3.18}$$

Y por otro lado,

$$\begin{aligned}
\mathbb{E}\Big(\sum_{n\geq 1}\mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{s<T_n\leq t\}}\big|\mathcal{G}_s\Big) &= \sum_{n\geq 1}\mathbb{E}(\mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{s<T_n\leq t\}}|\mathcal{G}_s)\\
&= \sum_{n\geq 1}\mathbb{P}(Y_{T_n}=j, s<T_n\leq t|\mathcal{G}_s)\\
&= \sum_{i\neq j}\mathbb{P}(Y_t=j|Y_s=i)\mathbb{1}_{\{Y_s=i\}}\\
&= \sum_{i\neq j}\mathbb{1}_{\{Y_s=i\}}\Big(\delta_{ij}+\int_s^t\sum_{l\in\mathcal{E}} q^{lj}\mathbb{P}(Y_{r^-}=l|Y_s=i)dr\Big)\\
&= \int_s^t \sum_{i\neq j}\sum_{l\in\mathcal{E}}\mathbb{1}_{\{Y_s=i\}}q^{lj}\mathbb{P}(Y_{r^-}=l|Y_s=i)dr, \qquad (3.19)
\end{aligned}$$

en lo anterior se utilizó la ecuación de Kolmogorov y la propiedad de Markov.

Entonces, por (3.18) y (3.19), se tiene que

$$\begin{aligned}
&\mathbb{E}\Big(\sum_{n\geq 1}\mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{T_n\leq t\}}\big|\mathcal{G}_s\Big) - \mathbb{E}\Big(\int_0^t\sum_{i\neq j} q^{ij}\mathbb{1}_{\{Y_{r^-}=i\}}\big|\mathcal{G}_s\Big)\\
&= \sum_{n\geq 1}\mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{T_n\leq s\}} - \int_0^s\sum_{i\neq j} q^{ij}\mathbb{1}_{\{Y_{r^-}=i\}}dr = Z_s,
\end{aligned}$$

con lo que se concluye la prueba. ∎

Luego, sea $v(t,w,x,i)$ de clase $C^{1,2,2}([0,T],\mathbb{R}^+,\mathbb{R})$ una solución de la ecuación HJB con crecimiento cuadrático y $h\in\mathcal{A}$ una estrategia admisible.

Ahora, definimos las siguientes medidas aleatorias, μ,ρ:

$$\begin{aligned}
\mu([0,t]\times\{j\}) &= \sum_{n\geq 1}\mathbb{1}_{\{Y_{T_n}=j\}}\mathbb{1}_{\{T_n\leq t\}}, \quad j\in\mathcal{E},\ \ t\in[0,T],\\
\rho([0,t]\times\{j\}) &= \int_0^t\sum_{i\neq j} q^{ij}\mathbb{1}_{\{Y_{r^-}=i\}}dr, \quad j\in\mathcal{E},\ \ t\in[0,T],
\end{aligned}$$

Por otro lado, observemos que

$$\sum_{t\leq s\leq T}[v(s,W_s^h,X_s,Y_s)-v(s,W_s^h,X_s,Y_{s^-})]$$

$$
\begin{aligned}
&= \sum_{n\geq 1}[v(T_n, W^h_{T_n}, X_{T_n}, Y_{T_n}) - v(T_n, W^h_{T_n}, X_{T_n}, Y_{T_n^-})]\mathbb{1}_{\{t\leq T_n\leq T\}} \\
&= \sum_{n\geq 1}\sum_{j\in\mathcal{E}}[v(T_n, W^h_{T_n}, X_{T_n}, j) - v(T_n, W^h_{T_n}, X_{T_n}, Y_{T_n^-})]\mathbb{1}_{\{t\leq T_n\leq T\}}\mathbb{1}_{\{Y_{T_n}=j\}} \\
&= \int_t^T \sum_{j\in\mathcal{E}}[v(T_n, W^h_{T_n}, X_{T_n}, j) - v(T_n, W^h_{T_n}, X_{T_n}, Y_{T_n^-})]\mu(dr\times\{j\}),
\end{aligned}
$$

donde $\{T_n\}_{n\in\mathbb{N}}$ es la sucesión de tiempos de salto de la cadena Y.

Por lo anterior se cumple que

$$
\begin{aligned}
\sum_{t\leq s\leq T}&[v(s, W^h_s, X_s, Y_s) - v(s, W^h_s, X_s, Y_{s^-})] = \\
&\int_t^T \sum_{j\in\mathcal{E}}[v(r, W^h_r, X_r, j) - v(r, W^h_r, X_r, Y_{r^-})](\mu-\rho)(dr\times\{j\}) \\
&+ \int_t^T \sum_{j\in\mathcal{E}}[v(s, W^h_s, X_s, j) - v(s, W^h_s, X_s, Y_{s^-})]q^{Y_{s^-}j}ds.
\end{aligned} \tag{3.20}
$$

Observación 3.1.1. *Por la fórmula de Itô, para $s\geq t$ y h una estrategia admisible, se cumple que*

$$
\begin{aligned}
dv(t, W^h_t, X_t, Y_t) =& v_t(t, W^h_t, X_t, Y_t)dt + v_w(t, W^h_t, X_t, Y_t)dW_t + v_x(t, W^h_t, X_t, Y_t)dX_t \\
&+ \frac{1}{2}\big[v_{ww}(t, W^h_t, X_t, Y_t)d[W,W]_t + v_{xx}(t, W^h_t, X_t, Y_t)d[X,X]_t \\
&+ v_{xw}(t, W^h_t, X_t, Y_t)d[W,X]_t\big] + \sum_{0\leq s\leq t}[v(s, W^h_s, X_s, Y_s) - v(s, W^h_s, X_s, Y_{s^-})],
\end{aligned}
$$

esto es,

$$
\begin{aligned}
&dv(s, W^h_s, X_s, Y_s) = \\
&\big(v_s(s, W^h_s, X_s, Y_s) + v_x(s, W^h_s, X_s, Y_s)\big[\Gamma_1 - \lambda^i_1(x-\alpha^i_1) - \lambda^i_2(x-\alpha^i_2)\big] \\
&+ W^h_s v_w(s, W^h_s, X_s, Y_s)\big[r + (h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)\mu_m + h^{(2)}\lambda^i_2(x-\alpha^i_2) - h^{(1)}\lambda^i_1(x-\alpha^i_1)\big] \\
&+ \frac{1}{2}W^2_s v_{ww}(s, W^h_s, X_s, Y_s)\big[\sigma^2_m(h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)^2 + \sigma^2(h^{(1)} + h^{(2)})^2 + (b_1h^{(1)})^2 + (b_2h^{(2)})^2\big] \\
&+ W^h_s v_{wx}(s, W^h_s, X_s, Y_s)\big[\sigma^2_m(\beta_1-\beta_2)(h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2) + b^2_1h^{(1)} - b^2_2h^{(2)}\big] \\
&+ \frac{1}{2}v_{xx}(s, W^h_s, X_s, Y_s)\big[(\beta_1-\beta_2)^2\sigma^2_m + b^2_1 + b^2_2\big]\big)ds + \sum_{0\leq s\leq s}[v(s, W^h_s, X_s, Y_s) - v(s, W^h_s, X_s, Y_{s^-})] \\
&+ \big[v_x(s, W^h_s, X_s, Y_s)(\beta_1-\beta_2)\sigma_m + \sigma_m W^h_s v_w(s, W^h_s, X_s, Y_s)(h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)\big]dB^{(m)}_s \\
&+ \big[b_1v_x(s, W^h_s, X_s, Y_s) + b_1h^{(1)}W^h_s v_w(s, W^h_s, X_s, Y_s)\big]dB^{(1)}_s \\
&+ \big[b_2h^{(2)}W^h_s v_w(s, W^h_s, X_s, Y_s) - b_2v_x(s, W^h_s, X_s, Y_s)\big]dB^{(2)}_s \\
&+ \big(W^h_s v_w(s, W^h_s, X_s, Y_s)\sigma(h^{(1)} + h^{(2)})\big)dB^{(0)}_s.
\end{aligned}
$$

De forma integral, por la ecuación (3.20), se sigue que

$$
\begin{aligned}
v(s, W_s^h, X_s, Y_s) =& v(t,w,x,i) + \int_t^t \big(v_t(t,w,x,i) + \mathcal{L}_{\mathbb{G}}^h v(r, W_r^h, X_r, Y_r)\big) dr \\
&+ \int_t^s \Big(v_x(r, W_r^h, X_r, Y_r)(\beta_1 - \beta_2)\sigma_m\Big) dB_r^{(m)} \\
&+ \int_t^s \Big(\sigma_m W_r^h v_w(r, W_r^h, X_r, Y_r)(h_r^{(m)} + h_r^{(1)}\beta_1 + h_r^{(2)}\beta_2)\Big) dB_r^{(m)} \\
&+ \int_t^s \Big(b_1 v_x(r, W_r^h, X_r, Y_r) + b_1 h_r^{(1)} W_r^h v_w(r, W_r^h, X_r, Y_r)\Big) dB_r^{(1)} \\
&+ \int_t^s \Big(b_2 h_r^{(2)} W_r^h v_w(r, W_r^h, X_r, Y_r) - b_2 v_x(r, W_r^h, X_r, Y_r)\Big) dB_r^{(2)} \\
&+ \int_t^s \Big(W_r^h v_w(r, W_r^h, X_r, Y_r)\sigma(h_r^{(1)} + h_r^{(2)})\Big) dB_r^{(0)} \\
&+ \int_t^s \Big(\sum_{j\in\mathcal{E}} [v(r, W_r^h, X_r, j) - v(r, W_r^h, X_r, Y_{r-})]\Big)(\mu - \rho)(dr \times \{j\}),
\end{aligned}
$$

donde $\mathcal{L}_{\mathbb{G}}^h$ denota el generador del proceso (W, X, Y).

Para $n \geq 1$ introducimos los siguientes tiempos de paro

$$\tau_n \coloneqq T \wedge \operatorname{ínf}\{s > t : \|(W_s, X_s, Y_s) - (w, x, i)\| \geq n\}.$$

Notemos que τ_n converge a T $\mathbb{P}$-c.s. Además, por la Observación 3.1.1, se satisface que

$$
\begin{aligned}
v(\tau_n, W_{\tau_n}^h, X_{\tau_n}, Y_{\tau_n}) =& v(t,w,x,i) + \int_t^{\tau_n} \big(v_t(t,w,x,i) + \mathcal{L}_{\mathbb{G}}^h v(r, W_r^h, X_r, Y_r)\big) dr \\
&+ \int_t^{\tau_n} \Big(v_x(r, W_r^h, X_r, Y_r)(\beta_1 - \beta_2)\sigma_m\Big) dB_r^{(m)} \\
&+ \int_t^{\tau_n} \Big(\sigma_m W_r^h v_w(r, W_r^h, X_r, Y_r)(h_r^{(m)} + h_r^{(1)}\beta_1 + h_r^{(2)}\beta_2)\Big) dB_r^{(m)} \\
&+ \int_t^{\tau_n} \Big(b_1 v_x(r, W_r^h, X_r, Y_r) + b_1 h_r^{(1)} W_r^h v_w(r, W_r^h, X_r, Y_r)\Big) dB_r^{(1)} \\
&+ \int_t^{\tau_n} \Big(b_2 h_r^{(2)} W_r^h v_w(r, W_r^h, X_r, Y_r) - b_2 v_x(r, W_r^h, X_r, Y_r)\Big) dB_r^{(2)} \\
&+ \int_t^{\tau_n} \Big(W_r^h v_w(r, W_r^h, X_r, Y_r)\sigma(h_r^{(1)} + h_r^{(2)})\Big) dB_r^{(0)} \\
&+ \int_t^{\tau_n} \Big(\sum_{j\in\mathcal{E}} [v(r, W_r^h, X_r, j) - v(r, W_r^h, X_r, Y_{r-})]\Big)(\mu - \rho)(dr \times \{j\}).
\end{aligned}
\tag{3.21}
$$

Luego, como v satisface la ecuación HJB, entonces de (3.21)

$$\begin{aligned}v(\tau_n, W^h_{\tau_n}, X_{\tau_n}, Y_{\tau_n}) \leq & v(t,w,x,i) + \int_t^{\tau_n} \Big(v_x(r, W^h_r, X_r, Y_r)(\beta_1 - \beta_2)\sigma_m \Big) dB^{(m)}_r \\ & + \int_t^{\tau_n} \Big(\sigma_m W^h_r v_w(r, W^h_r, X_r, Y_r)(h^{(m)}_r + h^{(1)}_r \beta_1 + h^{(2)}_r \beta_2) \Big) dB^{(m)}_r \\ & + \int_t^{\tau_n} \Big(b_1 v_x(r, W^h_r, X_r, Y_r) + b_1 h^{(1)}_r W^h_r v_w(r, W^h_r, X_r, Y_r) \Big) dB^{(1)}_r \\ & + \int_t^{\tau_n} \Big(b_2 h^{(2)}_r W^h_r v_w(r, W^h_r, X_r, Y_r) - b_2 v_x(r, W^h_r, X_r, Y_r) \Big) dB^{(2)}_r \\ & + \int_t^{\tau_n} \Big(W^h_r v_w(r, W^h_r, X_r, Y_r)\sigma(h^{(1)}_r + h^{(2)}_r) \Big) dB^{(0)}_r \\ & + \int_t^{\tau_n} \Big(\sum_{j\in\mathcal{E}} [v(r, W^h_r, X_r, j) - v(r, W^h_r, X_r, Y_{r^-})] \Big) (\mu - \rho)(dr \times \{j\}).\end{aligned} \tag{3.22}$$

Observemos que los integrandos respecto a $B^{(0)}, B^{(1)}, B^{(2)}$ y $B^{(m)}$ son acotados por definición de τ_n, ya que h cumple (3.1) y $v(t,w,x,i) \in C^{1,2,2}([0,T], \mathbb{R}^+, \mathbb{R})$. De lo anterior se sigue que tales integrales tienen esperanza cero.

Además, por definición de τ_n y al ser $v(t,w,x,i)$ de clase $C^{1,2,2}([0,T], \mathbb{R}^+, \mathbb{R})$ se tiene que el integrando

$$\sum_{j\in\mathcal{E}} [v(r, W^h_r, X_r, j) - v(r, W^h_r, X_r, Y_{r^-})]$$

es acotado. Por lo tanto, por el Lema 3.1.1 y el Lema 3.1 en [5], se sigue que la integral

$$\int_t^{\tau_n} \sum_{j\in\mathcal{E}} [v(r, W^h_r, X_r, j) - v(r, W^h_r, X_r, Y_{r^-})](\mu - \rho)(dr \times \{j\})$$

tiene esperanza cero.

Entonces, tomando esperanza en (3.22) obtenemos $\mathbb{E}_{t,w,x,i}[v(\tau_n, W^h_{\tau_n}, X_{\tau_n}, Y_{\tau_n})] \leq v(t,w,x,i)$. Además, por la continuidad de v, $\mathbb{P}$-c.s. se tiene que

$$\mathbb{E}_{t,w,x,i}[v(T, W_T, X_T, Y_{T^-})] = \mathbb{E}_{t,w,x,i}[\lim_{n\to\infty} v(\tau_n, W_{\tau_n}, X_{\tau_n}, Y_{\tau_n})].$$

Por otro lado, como v tiene crecimiento cuadrático,

$$|v(\tau_n, W_{\tau_n}, X_{\tau_n}, Y_{\tau_n})| \leq C\big(1 + |W_{\tau_n}|^2 + |X_{\tau_n}|^2 + |Y_{\tau_n}|^2\big) \leq C\big(1 + \sup_{s\leq T} \big\{|W_s|^2 + |X_s|^2 + |Y_s|^2\big\}\big),$$

y utilizando el Teorema 1.1.6, al ser las funciones $\lambda(\cdot)$ y $\alpha(\cdot)$ localmente Lipschitz continuas y con crecimiento lineal, se sigue que $\sup_{s\leq T} \big\{|W_s|^2 + |X_s|^2 + |Y_s|^2\big\}$ es integrable. Así, por convergencia dominada se satisface que

$$\mathbb{E}_{t,w,x,i}[v(T, W_T, X_T, Y_{T-})] = \lim_{n\to\infty} \mathbb{E}_{t,w,x,i}[v(\tau_n, W_{\tau_n}, X_{\tau_n}, Y_{\tau_n})] \leq v(t, w, x, i).$$

Luego, por la condición terminal $v(T, \cdot) = \log(\cdot)$, se cumple que $\mathbb{E}_{t,w,x,i}[\log(W_T^h)] \leq v(t, w, x, i)$. Finalmente, al tomar supremo sobre las estrategias admisibles, concluimos que $V(t, w, x, i) \leq v(t, w, x, i)$.

Para obtener la otra desigualdad, notemos que si $h_* = \left(h_*^{(m)}, h_*^{(1)}, h_*^{(2)}\right)$ es la estrategia óptima de la ecuación HJB entonces, para toda t,

$$v_t(t, w, x, i) + \mathcal{L}_{\mathbb{G}}^{h_*} v(t, w, x, i) = v_t(t, w, x, i) + \sup_{h\in\mathcal{A}} \mathcal{L}_{\mathbb{G}}^{h} v(t, w, x, i) = 0.$$

Así, de (3.21), bajo el mismo argumento anterior y la condición terminal $v(T, \cdot) = \log(\cdot)$, al tomar esperanza se satisface

$$\mathbb{E}_{t,w,x,i}[\log(W_T^h)] = \mathbb{E}_{t,w,x,i}[v(T, W_T^h, X_T, Y_T)] = v(t, w, x, i),$$

y así,

$$V(t, w, x, i) = \sup_{h\in\mathcal{A}} \mathbb{E}_{t,w,x,i}[\log(W_T^h)] \geq \mathbb{E}_{t,w,x,i}[\log(W_T^h)] = v(t, w, x, i).$$

De donde se sigue el resultado. ■

Observaciones respecto a la estrategia óptima

1. Notemos que $h_*^{(1)}$ y $h_*^{(2)}$ no dependen de los parámetros de mercado, β_1, β_2, σ_m y μ_m, esto se debe a que la exposición al riesgo de mercado de cada uno de los activos es cubierta al invertir en el índice de mercado, $S^{(m)}$.

2. Los parámetros ϱ_1 y ϱ_2 representan la variación relativa del riesgo idiosincrático, esto es, el papel de ϱ_1 se puede interpretar como un escalamiento de la contribución de la variación idiosincrática de $S^{(2)}$ y su *alfa* correspondiente, en $h_*^{(1)}$. De forma análoga para ϱ_2.

3. A pesar de que el índice de mercado, $S^{(m)}$, es independiente de la cadena de Markov Y, la estrategia asociada $h^{(m)}$ resulta ser *Markov modulada*, es decir, depende de los parámetros de cambio de régimen, λ_i, α_i, $i = 1, 2$.

4. Si consideramos $\beta_1 = \beta_2 = \beta$, $\alpha_1 = \alpha_2 = 0$, $\lambda_1^i = \lambda_1$ y $\lambda_2^i = \lambda_2$, para $i \in \{1, \ldots, k\}$, obtenemos la estrategia óptima bajo el modelo propuesto en [8] con función de utilidad

logarítmica. Explícitamente:

$$
\begin{aligned}
h_*^{(1)}(t,x) &= -\frac{x(\lambda_1+\varrho\lambda_2)}{b^2(1+\varrho)},\\
h_*^{(2)}(t,x) &= \frac{x(\lambda_2+\varrho\lambda_1)}{b^2(1+\varrho)},\\
h_*^{(m)}(t,x) &= \frac{\mu_m}{\sigma_m^2} - \beta\big(h_*^{(1)}(t,x)+h_*^{(2)}(t,x)\big),
\end{aligned}
$$

con $\varrho = \frac{\sigma^2}{\sigma^2+b^2}$.

3.2. Caso beta neutral

En esta sección se estudia el problema de optimización con la restricción de que la cartera de inversión es *beta neutral*. Más precisamente:

Definición 3.2.1. *Una estrategia $h^\beta = (h^{(\beta,m)}, h^{(\beta,1)}, h^{(\beta,2)})$ es beta neutral $\mathbb{G}$-admisible si es una estrategia autofinanciada y $\mathbb{G}$-predecible tal que*

$$\beta_1 h^{(\beta,1)} + \beta_2 h^{(\beta,2)} = 0, \qquad t\in[0,T], \tag{3.23}$$

y satisface

$$\mathbb{E}\bigg(\int_0^T (h_t^{(\beta,m)})^2 + (h_t^{(\beta,1)})^2 dt\bigg) < \infty. \tag{3.24}$$

Denotaremos por $\mathcal{A}^\beta$ al conjunto de estrategias beta neutrales $\mathbb{G}$-admisibles.

Recordemos que una estrategia se dice que es *beta nuetral* si la beta asociada es cero. Para este caso, notemos que la beta del portafolio es $\beta_1 h^{(\beta,1)} + \beta_2 h^{(\beta,2)}$, que justamente es la condición (3.23).

El valor del portafolio W^{h^β} asociado a la estrategia h^β seguirá la siguiente dinámica

$$dW_t^{h^\beta} = W_t^{h^\beta}\bigg[h_t^{(\beta,1)}\frac{dS_t^{(1)}}{S_t^{(1)}} + h_t^{(\beta,2)}\frac{dS_t^{(2)}}{S_t^{(2)}} + h_t^{(\beta,m)}\frac{dS_t^{(m)}}{S_t^{(m)}}\bigg] + rW_t^{h^\beta}\Big[1 - h_t^{(\beta,1)} - h_t^{(\beta,2)} - h_t^{(\beta,m)}\Big]dt,$$

que representa el capital ganado debido a cambios en los activos $S^{(m)}, S^{(1)}$ y $S^{(2)}$, más los intereses generados. Utilizando la restricción de neutralidad (3.23) y sustituyendo (2.1), (2.2) y (2.3), obtenemos

$$dW_t^{h^\beta} = W_t^{h^\beta}\Big[\big(r + h_t^{(\beta,m)}\mu_m + h_t^{(\beta,2)}\lambda_2(Y_t)(X_t-\alpha_2(Y_t)) - h_t^{(\beta,1)}\lambda_1(Y_t)(X_t-\alpha_1(Y_t))\big)dt$$

$$
\begin{aligned}
&+\sigma\big(h_t^{(\beta,1)}+h_t^{(\beta,2)}\big)dB_t^{(0)}+b_1h_t^{(\beta,1)}dB_t^{(1)}+b_2h_t^{(\beta,2)}dB_t^{(2)}+\sigma_m h_t^{(\beta,m)}dB_t^{(m)}\Big]\\
=&f^{\beta}(t,X_t,Y_t)W_t^{h^\beta}dt+\sum_{l=0}^{3}g_l^{\beta}(t)W_t^{h^\beta}dB_t^l,
\end{aligned}
$$

con $(B^0,B^1,B^2,B^3)\coloneqq(B^{(0)},B^{(1)},B^{(2)},B^{(m)})$, $f^{\beta}(t,X_t,Y_t)\coloneqq r+h_t^{(\beta,m)}\mu_m+h_t^{(\beta,2)}\lambda_2(Y_t)(X_t-\alpha_2(Y_t))-h_t^{(\beta,1)}\lambda_1(Y_t)(X_t-\alpha_1(Y_t))$, y

$$
\begin{aligned}
g_0^{\beta}(t)&\coloneqq\sigma(h_t^{(\beta,1)}+h_t^{(\beta,2)}),\\
g_1^{\beta}(t)&\coloneqq b_1h_t^{(\beta,1)},\\
g_2^{\beta}(t)&\coloneqq b_2h_t^{(\beta,2)},\\
g_3^{\beta}(t)&\coloneqq \sigma_m h_t^{(\beta,m)}.
\end{aligned}
$$

Por el Lema 1.1.1, para $u\le t$ se tiene que, $W_t^{h^\beta}$ está dado por

$$
\begin{aligned}
W_t^{h^\beta}&=W_u^{h^\beta}\exp\left(\int_u^t\left[f^{\beta}(s,X_s,Y_s)-\frac{1}{2}\sum_{l=0}^{3}(g_l^{\beta}(s))^2\right]ds+\sum_{l=0}^{3}\int_u^t g_l^{\beta}(s)dB_s^l\right)\\
&=:W_u^h\exp(G^{\beta}(u,t,X_s,Y_s;u\le s\le t)),
\end{aligned}
\tag{3.25}
$$

donde

$$
G^{\beta}(u,t,X_s,Y_s;u\le s\le t)=\int_u^t\left[f^{\beta}(s,X_s,Y_s)-\frac{1}{2}\sum_{l=0}^{3}(g_l^{\beta}(s))^2\right]ds+\sum_{l=0}^{3}\int_u^t g_l^{\beta}(s)dB_s^l.
$$

En el siguiente teorema se da la expresión de la estrategia beta neutral óptima y la función de valor del correspondiente problema de control.

Teorema 3.2.1. *Consideremos una inversión con función de utilidad logarítmica. Entonces la estrategia beta neutral óptima* $h_*^{\beta}=(h_*^{(\beta,m)},h_*^{(\beta,1)},h_*^{(\beta,2)})\in\mathcal{A}^{\beta}$ *está dada por*

$$
\begin{aligned}
h_*^{(\beta,1)}(t,x,i)&=-\frac{\lambda_1^i(x-\alpha_1^i)+\frac{\beta_1}{\beta_2}\lambda_2^i(x-\alpha_2^i)}{b_1^2+\frac{\beta_1^2}{\beta_2^2}b_2^2+\sigma^2\big[1-\frac{\beta_1}{\beta_2}\big]^2},\\
h_*^{(\beta,2)}(t,x,i)&=-\frac{\beta_1}{\beta_2}h_*^{(\beta,1)}(t,x,i),\\
h_*^{(\beta,m)}(t,x,i)&=\frac{\mu_m}{\sigma_m^2}.
\end{aligned}
\tag{HBNO}
$$

Además, la función de valor es de la forma

$$
V(t,w,x,i)=\log(w)+m(t,i)x^2+n(t,i)x+u(t,i),
$$

donde las funciones $m(t,i)$, $n(t,i)$ y $u(t,i)$ para cada $i \in \mathcal{E}$ resuelven el siguiente sistema de ecuaciones diferenciales ordinarias

$$
\begin{aligned}
m_t(t,i) - 2(\lambda_1^i + \lambda_2^i)m(t,i) + \sum_{j\in\mathcal{E}} m(t,j)q^{ij} + \Phi_1^i = 0,\\
n_t(t,i) - (\lambda_1^i + \lambda_2^i)n(t,i) + \sum_{j\in\mathcal{E}} n(t,j)q^{ij} + 2(\Gamma_1 + \lambda_1^i\alpha_1^i + \lambda_2^i\alpha_2^i)m(t,i) - \Phi_2^i = 0,\\
u_t(t,i) + \sum_{j\in\mathcal{E}} u(t,j)q^{ij} + \Gamma_2 m(t,i) + (\Gamma_1 + \lambda_1^i\alpha_1^i + \lambda_2^i\alpha_2^i)n(t,i) + \Phi_3^i = 0,
\end{aligned}
$$

con condiciones terminales $m(T,\cdot) = 0$, $n(T,\cdot) = 0$ y $u(T,\cdot) = 0$, y los parámetros Γ_1, Γ_2, Φ_k^i están dados por

$$
\begin{aligned}
\Gamma_1 =& (\beta_1 - \beta_2)\mu_m - \frac{1}{2}((\beta_1^2 - \beta_2^2)\sigma_m^2 + b_1^2 - b_2^2),\\
\Gamma_2 =& \sigma_m^2(\beta_1 - \beta_2)^2 + b_1^2 + b_2^2,\\
\Phi_1^i =& \frac{(\beta_2\lambda_1^i + \beta_1\lambda_2^i)^2}{2\left[b_1^2\beta_2^2 + b_2^2\beta_1^2 + \sigma^2(\beta_1 - \beta_2)^2\right]},\\
\Phi_2^i =& \frac{(\alpha_1^i\beta_2\lambda_1^i + \alpha_2^i\beta_1\lambda_2^i)(\beta_1\lambda_2^i + \beta_2\lambda_1^i)}{b_1^2\beta_2^2 + b_2^2\beta_1^2 + \sigma^2(\beta_1 - \beta_2)^2},\\
\Phi_3^i =& r + \frac{\mu_m^2}{2\sigma_m^2} + \frac{(\alpha_1^i\beta_2\lambda_1^i + \alpha_2^i\beta_1\lambda_2^i)^2}{2\left[b_1^2\beta_2^2 + b_2^2\beta_1^2 + \sigma^2(\beta_1 - \beta_2)^2\right]}.
\end{aligned}
$$

En lo anterior m_t, n_t y u_t denotan las derivadas parciales respecto a t de las funciones m, n y u, respectivamente.

Demostración.

La demostración será análoga al Teorema 3.1.1. Esto es, a partir de la ecuación de Hamilton-Jacobi-Bellman obtener la estrategia beta neutral óptima $h_*^\beta = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)})$. Luego proponer un candidato a la función de valor y por último verificar que tal candidato coincide con la solución de la ecuación HJB.

1. Estrategia Óptima

Supongamos que la función de valor $V(\cdot,i)$ es de clase $C^{1,2,2}([0,T],\mathbb{R}^+,\mathbb{R})$, esto implica que satisface la ecuación de HJB dada, para cada $i \in \mathcal{E}$, por (3.4).

Por otro lado, de forma similar al Teorema 3.1.1, por (3.25)

$$
V(t,w,x,i) = \log(w) + \sup_{h^\beta\in\mathcal{A}^\beta} \mathbb{E}_{t,w,x,i}[G^\beta(t,T,X_s,Y_s; t \le s \le T)] =: \log(w) + \nu(t,x,i), \quad (3.26)
$$

para alguna función $\nu(t,x,i)$, con $\nu(T,\cdot) = 0$.

Sustituyendo (3.26) en (3.4) y usando la hipótesis en (3.23), obtenemos que

$$
\begin{aligned}
v_t(t,w,x,i) &+ \sup_{h^\beta\in\mathcal{A}^\beta} \mathcal{L}_{\mathbb{G}}^{h^\beta} V(t,w,x,i) \\
&= \nu_t(t,x,i) + \sup_{h^\beta\in\mathcal{A}^\beta} \Bigg[\nu_x(t,x,i)\big[\Gamma_1 - \lambda_1^i(x-\alpha_1^i) - \lambda_2^i(x-\alpha_2^i)\big] \\
&\qquad + r + h^{(\beta,m)}\mu_m + h^{(\beta,2)}\lambda_2^i(x-\alpha_2^i) - h^{(\beta,1)}\lambda_1^i(x-\alpha_1^i) \\
&\qquad - \frac{1}{2}\big[\sigma_m^2(h^{(\beta,m)})^2 + \sigma^2(h^{(\beta,1)}+h^{(\beta,2)})^2 + (b_1h^{(\beta,1)})^2 + (b_2h^{(\beta,2)})^2\big] \\
&\qquad + \frac{1}{2}\Gamma_2\nu_{xx}(t,x,i) + \sum_{j\in\mathcal{E}} \nu(t,x,j)q^{ij}\Bigg] \\
&= 0.
\end{aligned}
\tag{3.27}
$$

Para hallar el supremo que aparece en (3.27), al contar con la restricción de neutralidad dada por (3.23), se obtendrá la estrategia óptima $h_*^\beta = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)})$ mediante el método de multiplicadores de Lagrange. Primero encontrando puntos críticos con las condiciones de primer orden y luego verificando que éstos son máximos con las condiciones de segundo orden.

Por las condiciones de primer orden, considerando la restricción (3.23) y derivando respecto de $h_*^{(\beta,m)}, h_*^{(\beta,1)}$ y $h_*^{(\beta,2)}$, respectivamente de (3.27), se tiene que

$$
\frac{\mu_m}{\sigma_m^2} - {h_*}^{(\beta,m)} = 0, \tag{3.28}
$$

$$
-\lambda_1^i(x-\alpha_1^i) - \sigma^2(h_*^{(\beta,1)} + h_*^{(\beta,2)}) - b_1^2 h_*^{(\beta,1)} = \Lambda\beta_1, \tag{3.29}
$$

$$
\lambda_2^i(x-\alpha_2^i) - \sigma^2(h_*^{(\beta,1)} + h_*^{(\beta,2)}) - b_2^2 h_*^{(\beta,2)} = \Lambda\beta_2, \tag{3.30}
$$

por lo que

$$
h_*^{(\beta,m)} = \frac{\mu_m}{\sigma_m^2},
$$

además,

$$
\frac{1}{\beta_1}\big[-\lambda_1^i(x-\alpha_1^i) - h_*^{(\beta,1)}(\sigma^2+b_1^2) - \sigma^2 h_*^{(\beta,2)}\big] = \frac{1}{\beta_2}\big[\lambda_2^i(x-\alpha_2^i) - h_*^{(\beta,2)}(\sigma^2+b_2^2) - \sigma^2 h_*^{(\beta,1)}\big],
$$

esto es,

$$
h_*^{(\beta,2)}\left[\frac{\sigma^2+b_2^2}{\beta_2} - \frac{\sigma^2}{\beta_1}\right] = h_*^{(\beta,1)}\left[\frac{\sigma^2+b_1^2}{\beta_1} - \frac{\sigma^2}{\beta_2}\right] + \frac{\lambda_2^i}{\beta_2}(x-\alpha_2^i) + \frac{\lambda_1^i}{\beta_1}(x-\alpha_1^i).
$$

Luego, por (3.23), se tiene que

$$
-\beta_1 h_*^{(\beta,1)} = \frac{\beta_1\beta_2}{\beta_1(\sigma^2+b_2^2)-\sigma^2\beta_2}\left[\lambda_2^i(x-\alpha_2^i) + \frac{\beta_2}{\beta_1}\lambda_1^i(x-\alpha_1^i)\right] + h_*^{(\beta,1)}\left[\frac{\beta_2[\beta_2(\sigma^2+b_1^2)-\beta_1\sigma^2]}{\beta_1(\sigma^2+b_2^2)-\sigma^2\beta_2}\right],
$$

entonces,

$$h_*^{(\beta,1)}\frac{\sigma^2(\beta_1-\beta_2)^2+\beta_1^2b_2^2+\beta_2^2b_1^2}{\beta_1(\sigma^2+b_2^2)-\beta_2\sigma^2}=-\frac{\beta_1\beta_2}{\beta_1(\sigma^2+b_2^2)-\sigma^2\beta_2}\left[\lambda_2^i(x-\alpha_2^i)+\frac{\beta_2}{\beta_1}\lambda_1^i(x-\alpha_1^i)\right].$$

Por lo tanto,

$$h_*^{(\beta,1)}\left[\sigma^2\left(1-\frac{\beta_1}{\beta_2}\right)^2+b_2^2\left(\frac{\beta_1}{\beta_2}\right)^2+b_1^2\right]=-\left[\frac{\beta_1}{\beta_2}\lambda_2^i(x-\alpha_2^i)+\lambda_1^i(x-\alpha_1^i)\right],$$

de donde se concluye que

$$h_*^{(\beta,1)}=-\frac{\lambda_2^i(x-\alpha_2^i)\frac{\beta_1}{\beta_2}+\lambda_1^i(x-\alpha_1^i)}{\sigma^2\big(1-\frac{\beta_1}{\beta_2}\big)^2+b_2^2\big(\frac{\beta_1}{\beta_2}\big)^2+b_1^2}$$

y

$$h_*^{(\beta,2)}=-\frac{\beta_1}{\beta_2}h_*^{(\beta,1)}.$$

Para las condiciones de segundo orden, obteniendo las segundas derivadas se sigue que

$$A^\beta=(a_{ij}^\beta)_{\{i,j=1,2,3\}}:=\begin{pmatrix}-\sigma_m^2 & 0 & 0\\ 0 & -(\sigma^2+b_1^2) & -\sigma^2\\ 0 & -\sigma^2 & -(\sigma^2+b_2^2)\end{pmatrix}.$$

Veamos ahora que la matriz A^β es definida no positiva, equivalentemente, que se satisface

$$a_{11}^\beta<0,$$
$$\frac{1}{a_{11}^\beta}\begin{vmatrix}a_{11}^\beta & a_{12}^\beta\\ a_{12}^\beta & a_{22}^\beta\end{vmatrix}<0,$$
$$\frac{\det(A^\beta)}{\begin{vmatrix}a_{11}^\beta & a_{12}^\beta\\ a_{12}^\beta & a_{22}^\beta\end{vmatrix}}<0,$$

donde $\begin{vmatrix}a_{11}^\beta & a_{12}^\beta\\ a_{12}^\beta & a_{22}^\beta\end{vmatrix}$ denota el menor de la matriz A^β.

Notemos que, $\det(A^\beta)=-\sigma_m^2[b_1^2b_2^2+\sigma^2(b_1^2+b_2^2)]<0$ y

$$\begin{vmatrix}a_{11}^\beta & a_{12}^\beta\\ a_{12}^\beta & a_{22}^\beta\end{vmatrix}=\sigma_m^2(\sigma_1^2+b_1^2)>0,$$

de donde se verifica que $h_*^\beta = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)})$ en efecto está dada por (HBNO).

2. Existencia

De la ecuación (3.29)

$$\Lambda = \frac{1}{\beta_1}\Big[-\lambda_1^i(x-\alpha_1^i) - \sigma^2(h_*^{(\beta,1)} + h_*^{(\beta,2)}) - b_1^2 h_*^{(\beta,1)}\Big].$$

Sustituyendo Λ en (3.30)

$$\frac{\beta_2}{\beta_1}\Big[-\lambda_1^i(x-\alpha_1^i) - \sigma^2(h_*^{(\beta,1)} + h_*^{(\beta,2)}) - b_1^2 h_*^{(\beta,1)}\Big] = \lambda_2^i(x-\alpha_2^i) - \sigma^2(h_*^{(\beta,1)} + h_*^{(\beta,2)}) - b_2^2 h_*^{(\beta,2)},$$

multiplicando la ecuación anterior por $h_*^{(\beta,2)}$ y utilizando (3.23) se sigue que

$$-\lambda_1^i h_*^{(\beta,1)}(x-\alpha_1^i) + \lambda_2^i h_*^{(\beta,2)}(x-\alpha_2^i) - \sigma^2(h_*^{(\beta,1)} + h_*^{(\beta,2)})^2 - b_1^2(h_*^{(\beta,1)})^2 - b_2^2(h_*^{(\beta,2)})^2 = 0. \tag{3.31}$$

Por otro lado, sustituyendo $h_*^\beta = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)})$ en (3.27)

$$\begin{aligned}
&\nu_t(t,x,i) + \nu_x(t,x,i)\big[\Gamma_1 - \lambda_1^i(x-\alpha_1^i) - \lambda_2^i(x-\alpha_2^i)\big] + r + \frac{\mu_m^2}{\sigma_m^2} \\
&\quad - \frac{1}{2}\left(\frac{\mu_m^2}{\sigma_m^2} + \sigma^2(h_*^{(\beta,1)} + h_*^{(\beta,2)})^2 + (b_1 h_*^{(\beta,1)})^2 + (b_2 h_*^{(\beta,2)})^2\right) \\
&\quad + h_*^{(\beta,2)}\lambda_2^i(x-\alpha_2^i) - h_*^{(\beta,1)}\lambda_1^i(x-\alpha_1^i) + \frac{1}{2}\Gamma_2 \nu_{xx}(t,x,i) + \sum_{j\in\mathcal{E}} \nu(t,x,j)q^{ij} = 0,
\end{aligned}$$

y por (3.31)

$$\begin{aligned}
&\nu_t(t,x,i) + \nu_x(t,x,i)\big[\Gamma_1 - \lambda_1^i(x-\alpha_1^i) - \lambda_2^i(x-\alpha_2^i)\big] + r + \frac{\mu_m^2}{2\sigma_m^2} \\
&\quad + \frac{1}{2}\big(h_*^{(\beta,2)}\lambda_2^i(x-\alpha_2^i) - h_*^{(\beta,1)}\lambda_1^i(x-\alpha_1^i)\big) + \frac{1}{2}\Gamma_2 \nu_{xx}(t,x,i) + \sum_{j\in\mathcal{E}} \nu(t,x,j)q^{ij} = 0.
\end{aligned}$$

Pero observemos que

$$-h_*^{(\beta,1)}\lambda_1^i(x-\alpha_1^i) = \frac{(\lambda_1^i)^2(x^2 - 2\alpha_1^i x + (\alpha_1^i)^2) + \frac{\beta_1}{\beta_2}\lambda_1^i\lambda_2^i(x^2 - x(\alpha_1^i + \alpha_2^i) + \alpha_1^i\alpha_2^i)}{b_1^2 + \frac{\beta_1^2}{\beta_2^2}b_2^2 + \sigma^2\big(1 - \frac{\beta_1}{\beta_2}\big)^2}$$

y

$$h_*^{(\beta,2)}\lambda_2^i(x-\alpha_2^i)=\frac{\beta_1\big[\lambda_1^i\lambda_2^i(x^2-x(\alpha_1^i\alpha_2^i)+\alpha_1^i\alpha_2^i)+\frac{\beta_1}{\beta_2}(\lambda_2^i)^2(x^2-2x\alpha_2^i+(\alpha_2^i)^2)\big]}{\beta_2\big[b_1^2+\frac{\beta_1^2}{\beta_2^2}b_2^2+\sigma^2\big(1-\frac{\beta_1}{\beta_2}\big)^2\big]},$$

por lo que,

$$\begin{aligned}\nu_t(t,x,i)&+\nu_x(t,x,i)\big[\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i-(\lambda_1^i+\lambda_2^i)x\big]+\Phi_1^i x^2-\Phi_2^i x+\Phi_3^i\\&+\frac{1}{2}\Gamma_2\nu_{xx}(t,x,i)+\sum_{j\in\mathcal{E}}\nu(t,x,j)q^{ij}=0,\end{aligned}\tag{3.32}$$

con Φ_1^i,Φ_2^i y Φ_3^i dadas como en el enunciado del teorema.

Luego, como función de valor proponemos $\nu(t,x,i)=m(t,i)x^2+n(t,i)x+u(t,i)$, la cual al sustituirla en (3.32), para cualquier $(t,x)\in[0,T]\times\mathbb{R}$, se satisface

$$\begin{aligned}&m_t(t,i)x^2+n_t(t,i)x+u_t(t,i)+\big[\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i-(\lambda_1^i+\lambda_2^i)x\big][2m(t,i)x+n(t,i)]\\&\quad+\Phi_1^i x^2-\Phi_2^i x+\Phi_3^i+\Gamma_2 m(t,i)+\sum_{j\in\mathcal{E}}q^{ij}[m(t,j)x^2+n(t,j)x+u(t,j)]\\&=x^2\left[m_t(t,i)-2m(t,i)(\lambda_1^i+\lambda_2^i)+\Phi_1^i+\sum_{j\in\mathcal{E}}q^{ij}m(t,j)\right]\\&\quad+x\left[n_t(t,i)+2m(t,i)\big[\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i\big]-n(t,i)(\lambda_1^i+\lambda_2^i)-\Phi_2^i+\sum_{j\in\mathcal{E}}q^{ij}n(t,j)\right]\\&\quad+u_t(t,i)+n(t,i)[\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i]+\Phi_3^i+\Gamma_2 m(t,i)+\sum_{j\in\mathcal{E}}q^{ij}u(t,j)=0.\end{aligned}$$

Entonces, si para cada $i\in\mathcal{E}$, las funciones $m(t,i)$, $n(t,i)$ y $u(t,i)$ resuelven el siguiente sistema de ecuaciones diferenciales ordinarias

$$\begin{aligned}m_t(t,i)-2(\lambda_1^i+\lambda_2^i)m(t,i)+\sum_{j\in\mathcal{E}}m(t,j)q^{ij}+\Phi_1^i&=0,\\n_t(t,i)-(\lambda_1^i+\lambda_2^i)n(t,i)+\sum_{j\in\mathcal{E}}n(t,j)q^{ij}+2(\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i)m(t,i)-\Phi_2^i&=0,\\u_t(t,i)+\sum_{j\in\mathcal{E}}u(t,j)q^{ij}+\Gamma_2 m(t,i)+(\Gamma_1+\lambda_1^i\alpha_1^i+\lambda_2^i\alpha_2^i)n(t,i)+\Phi_3^i&=0,\end{aligned}$$

con condiciones terminales $m(T,\cdot)=0, n(T,\cdot)=0$ y $u(T,\cdot)=0$, se cumple que la función $\nu(t,x,i)=m(t,i)x^2+n(t,i)x+u(t,i)$ resuelve la ecuación de HJB. Por lo tanto, la función de valor está dada por $V(t,w,x,i)=\log(w)+\nu(t,x,i)=\log(w)+m(t,i)x^2+n(t,i)x+u(t,i)$.

3. Verificación.

Consideremos $v(t,w,x,i)$ de clase $C^{1,2,2}([0,T],\mathbb{R}^+,\mathbb{R})$ y con crecimiento cuadrático una solución de la ecuación HJB y $h^\beta \in \mathcal{A}^\beta$ una estrategia admisible.

Por la Observación 3.1.1

$$\begin{aligned}
v(T, W_T^{h^\beta}, X_T, Y_T) =& v(t,w,x,i) + \int_t^T \big(v_t(t,w,x,i) + \mathcal{L}_{\mathbb{G}}^{h^\beta} v(r, W_r^{h^\beta}, X_r, Y_r)\big)dr \\
&+ \int_t^T \Big(v_x(r, W_r^{h^\beta}, X_r, Y_r)(\beta_1 - \beta_2)\sigma_m\Big) dB_r^{(m)} \\
&+ \int_t^T \Big(b_1 v_x(r, W_r^{h^\beta}, X_r, Y_r) + b_1 h_r^{(\beta,1)} W_r^{h^\beta} v_w(r, W_r^{h^\beta}, X_r, Y_r)\Big) dB_r^{(1)} \\
&+ \int_t^T \Big(b_2 h_r^{(\beta,2)} W_r^{h^\beta} v_w(r, W_r^{h^\beta}, X_r, Y_r) - b_2 v_x(r, W_r^{h^\beta}, X_r, Y_r)\Big) dB_r^{(2)} \\
&+ \int_t^T \Big(W_r^{h^\beta} v_w(r, W_r^{h^\beta}, X_r, Y_r)\sigma(h_r^{(\beta,1)} + h_r^{(\beta,2)})\Big) dB_r^{(0)} \\
&+ \int_t^T \Big(\sigma_m W_r^{h^\beta} v_w(r, W_r^{h^\beta}, X_r, Y_r) h_r^{(\beta,m)}\Big) dB_r^{(m)} \\
&+ \int_t^T \Big(\sum_{j\in\mathcal{E}} [v(r, W_r^{h^\beta}, X_r, j) - v(r, W_r^{h^\beta}, X_r, Y_{r^-})]\Big)(\mu - \rho)(dr \times \{j\}).
\end{aligned} \tag{3.33}$$

Como v satisface la ecuación HJB, entonces se sigue que

$$\begin{aligned}
v(T, W_T^{h^\beta}, X_T, Y_T) \leq& v(t,w,x,i) + \int_t^T \Big(v_x(r, W_r^{h^\beta}, X_r, Y_r)(\beta_1 - \beta_2)\sigma_m\Big) dB_r^{(m)} \\
&+ \int_t^T \Big(b_1 v_x(r, W_r^{h^\beta}, X_r, Y_r) + b_1 h_r^{(\beta,1)} W_r^{h^\beta} v_w(r, W_r^{h^\beta}, X_r, Y_r)\Big) dB_r^{(1)} \\
&+ \int_t^T \Big(b_2 h_r^{(\beta,2)} W_r^{h^\beta} v_w(r, W_r^{h^\beta}, X_r, Y_r) - b_2 v_x(r, W_r^{h^\beta}, X_r, Y_r)\Big) dB_r^{(2)} \\
&+ \int_t^T \Big(W_r^{h^\beta} v_w(r, W_r^{h^\beta}, X_r, Y_r)\sigma(h_r^{(\beta,1)} + h_r^{(\beta,2)})\Big) dB_r^{(0)} \\
&+ \int_t^T \Big(\sigma_m W_r^{h^\beta} v_w(r, W_r^{h^\beta}, X_r, Y_r) h_r^{(\beta,m)}\Big) dB_r^{(m)} \\
&+ \int_t^T \Big(\sum_{j\in\mathcal{E}} [v(r, W_r^{h^\beta}, X_r, j) - v(r, W_r^{h^\beta}, X_r, Y_{r^-})]\Big)(\mu - \rho)(dr \times \{j\}).
\end{aligned} \tag{3.34}$$

Bajo el mismo argumento del Teorema 3.1.1, localizando se puede ver que las integrales respecto a $B^{(0)}$, $B^{(1)}$, $B^{(2)}$, $B^{(m)}$ y $(\mu - \rho)$ tienen esperanza cero. Entonces, al tomar esperanza en (3.34) y usando la condición terminal $v(T,\cdot) = \log(\cdot)$, se cumple que $\mathbb{E}_{t,w,x,i}[\log(W_T^{h^\beta})] \leq v(t,w,x,i)$. Finalmente, al tomar supremo sobre las estrategias admisibles se concluye que $V(t,w,x,i) \leq v(t,w,x,i)$.

Por otro lado, si $h_*^\beta = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)})$ es la estrategia óptima de la ecuación HJB, entonces

$$v_t(t,w,x,i) + \mathcal{L}_{\mathbb{G}}^{h_*^\beta} V(t,w,x,i) = v_t(t,w,x,i) + \sup_{h^\beta \in \mathcal{A}^\beta} \mathcal{L}_{\mathbb{G}}^{h^\beta} V(t,w,x,i) = 0.$$

Luego, por (3.33) y la condición terminal $v(T,\cdot) = \log(\cdot)$, se satisface que

$$\mathbb{E}_{t,w,x,i}[\log(W_T^{h^\beta})] = \mathbb{E}_{t,w,x,i}[v(T, W_T^{h^\beta}, X_T, Y_T)] = v(t,w,x,i),$$

por lo que

$$V(t,w,x,i) = \sup_{h \in \mathcal{A}^\beta} \mathbb{E}_{t,w,x,i}[\log(W_T^{h^\beta})] \geq \mathbb{E}_{t,w,x,i}[\log(W_T^{h^\beta})] = v(t,w,x,i).$$

■

Observaciones respecto a la estrategia óptima

1. Notemos que el cociente $\frac{\beta_1}{\beta_2}$ tiene la misma interpretación que ϱ_2 en el Teorema 3.1.1.
2. Si consideramos $\beta_1 = \beta_2$ obtenemos una estrategia *delta neutra*, es decir, aquella en donde se invierte la misma cantidad de dinero en ambos activos, $h_*^{(1)} = -h_*^{(2)}$. Más precisamente:

$$h_*^{(1)}(t,x,i) = -h_*^{(2)}(t,x,i) = -\frac{\lambda_1^i(x-\alpha_1^i) + \lambda_2^i(x-\alpha_2^i)}{b_1^2 + b_2^2},$$
$$h_*^{(m)}(t,x,i) = \frac{\mu_m}{\sigma_m^2}.$$

Capítulo 4

Optimización bajo información parcial

En el capítulo anterior se supone que la cadena de Markov que rige los *alfas* de los activos cointegrados es observable por el inversionista. Sin embargo, tales errores dependen de factores económicos o de mercado, que por su naturaleza son estocásticos y por ende resulta natural suponer que el inversionista no tenga acceso a este tipo de información [2]. De aquí la importancia de estudiar el problema de optimización desde la perspectiva de *información parcial*, lo cual constituye el objetivo de este capítulo.

Análogamente al capítulo anterior, se planteará el problema de control a resolver, se obtendrá la estrategia óptima y se caracterizará la función de valor como la única solución a un sistema de ecuaciones diferenciales parciales derivado de la ecuación de Hamilton-Jacobi-Bellman, tanto para el caso clásico como el caso *beta neutral*. Para ello, primero se resuelve el *problema de filtros*, lo cual permite pasar del problema de información parcial a uno equivalente de información completa [3].

4.1. Problema de filtros

Dado que el inversionista no tendrá acceso a toda la información, para poder resolver el problema de optimización, mediante filtros se reducirá el problema de control a uno equivalente de información completa. Esto es, se reemplazará la cadena de Markov Y, que no es observable, por su proyección sobre la filtración observada; a esto se le conoce como el *problema de filtros* [3].

El modelo que se empleará seguirá siendo el descrito en el Capítulo 2, recordemos que la filtración $\mathbb{G}$ representa todo el flujo de información, consideraremos ahora la filtración $\mathbb{F} := \{\mathcal{F}_t, t \in [0, T]\}$, con $\mathcal{F}_t = \sigma(S_r^{(m)}, S_r^{(1)}, S_r^{(2)}, 0 \leq r \leq t)$ y donde $S^{(m)}, S^{(1)}, S^{(2)}$ representarán el *benchmark* y los activos cointegrados, respectivamente. La información a la que tendrá acceso el inversionista será la generada por $\mathbb{F}$. Bajo este contexto, un filtro estará definido como sigue:

Definición 4.1.1. ***Filtro.*** *Para $f : \mathcal{E} \to \mathbb{R}$, definimos al filtro $\pi(f)$ como la proyección del proceso $f(Y_t)$ sobre la filtración $\mathcal{F}_t$, esto es, para $t \in [0, T]$*

$$\pi_t(f) := \mathbb{E}\big[f(Y_t)|\mathcal{F}_t\big] = \sum_{i \in \mathcal{E}} f(e_i)\mathbb{P}(Y_t = e_i|\mathcal{F}_t),$$

donde $\mathcal{E} = \{e_1, \ldots, e_k\}$ denota el espacio de estados de la cadena Y.

Notación 4.1.1. *Para el desarrollo de este capítulo se utilizará la siguiente notación*

$$\begin{aligned}
\pi_t^i &:= \mathbb{P}(Y_t = e_i|\mathcal{F}_t), \\
\boldsymbol{\pi} &:= (\pi^1, \ldots, \pi^k)^T, \\
\mu_1(X_t, Y_t) &:= -\lambda_1(Y_t)(X_t - \alpha_1(Y_t)), \\
\mu_2(X_t, Y_t) &:= \lambda_2(Y_t)(X_t - \alpha_2(Y_t)), \\
\boldsymbol{\mu}_i(X_t) &:= (\mu_i(X_t, e_1), \ldots, \mu_i(X_t, e_k)).
\end{aligned}$$

Introduciremos a continuación los procesos de *observación*, que están en función de Y y a partir de los cuales definiremos al proceso de *innovación* que nos permitirá hallar, para cada $i \in \mathcal{E}$, la dinámica del proceso $\pi_t^i = \mathbb{P}(Y_t = e_i|\mathcal{F}_t)$, para $t \in [0, T]$.

Definición 4.1.2. *Dada $t \in [0, T]$, definimos a los procesos de observación como*

$$R_t^{(1)} := \sigma B_t^{(0)} + b_1 B_t^{(1)} - \int_0^t \lambda_1(Y_s)(X_s - \alpha_1(Y_s))ds, \tag{4.1}$$

$$R_t^{(2)} := \sigma B_t^{(0)} + b_2 B_t^{(2)} + \int_0^t \lambda_2(Y_s)(X_s - \alpha_2(Y_s))ds, \tag{4.2}$$

y a los procesos estocásticos $Z^{(1)}$ y $Z^{(2)}$ por

$$Z_t^{(1)} := \frac{\sigma B_t^{(0)} + b_1 B_t^{(1)}}{\sigma_1}, \tag{4.3}$$

$$Z_t^{(2)} := \frac{\sigma B_t^{(0)} + b_2 B_t^{(2)}}{\sigma_2}, \tag{4.4}$$

con $\sigma_1 = \sqrt{\sigma^2 + b_1^2}$ y $\sigma_2 = \sqrt{\sigma^2 + b_2^2}$.

Observación 4.1.1. *De las ecuaciones* (2.2) *y* (2.3)*, por el Lema 1.1.1, podemos expresar a $S^{(1)}$ y $S^{(2)}$ como*

$$S_t^{(1)} = S_0^{(1)} \exp\bigg\{ \int_0^t \big(r + \beta_1\mu_m - \lambda_1(Y_s)[X_s - \alpha_1(Y_s)]\big)ds$$

$$
\begin{aligned}
&\qquad\qquad - \frac{t}{2}\big(\beta_1^2\sigma_m^2 + \sigma^2 + b_1^2\big) + \beta_1\sigma_m B_t^{(m)} + \sigma B_t^{(0)} + b_1 B_t^{(1)}\bigg\},\\
S_t^{(2)} &= S_0^{(2)} \exp\bigg\{\int_0^t \big(r + \beta_2\mu_m + \lambda_2(Y_s)[X_s - \alpha_2(Y_s)]\big)ds\\
&\qquad\qquad - \frac{t}{2}\big(\beta_2^2\sigma_m^2 + \sigma^2 + b_2^2\big) + \beta_2\sigma_m B_t^{(m)} + \sigma B_t^{(0)} + b_2 B_t^{(2)}\bigg\}.
\end{aligned}
$$

De donde obtenemos que

$$
R_t^{(1)} = \ln S_t^{(1)} - J_1(t, B^{(m)}) \quad y \quad R_t^{(2)} = \ln S_t^{(2)} - J_2(t, B^{(m)}),
$$

con $J_l(t, B^{(m)}) = \beta_l\sigma_m B_t^{(m)} + \ln S_0^{(l)} + t\big[r + \beta_l\mu_m - \frac{1}{2}(\beta_l^2\sigma_m^2 + \sigma^2 + b_l^2)\big]$, $l = 1, 2$.

Por lo tanto, al ser $R^{(1)}$ *y* $R^{(2)}$ *función de los procesos* $S^{(1)}$ *y* $S^{(2)}$, *la* σ*-álgebra generada por* $(S^{(1)}, S^{(2)}, S^{(m)})$ *y* $(R^{(1)}, R^{(2)}, S^{(m)})$ *coinciden.*

Con los siguientes resultados se construirá un proceso $\tilde{Z}^{(2)}$ independiente del proceso $Z^{(1)}$, para finalmente poder definir al proceso de *innovación* I.

Lema 4.1.1. *Los procesos* $Z^{(1)}$ *y* $Z^{(2)}$ *son* $(\mathbb{G}, \mathbb{P})$*-movimientos brownianos correlacionados con coeficiente de correlación* $\rho := \frac{\sigma^2}{\sigma_1\sigma_2}$.

Demostración. Notemos que $Z^{(1)}$ y $Z^{(2)}$ son $(\mathbb{G}, \mathbb{P})$-movimientos brownianos, pues $(B^{(0)}, B^{(1)}, B^{(2)})$ es $(\mathbb{G}, \mathbb{P})$-movimiento browniano tridimensional.

Por otro lado, recordemos que la correlación de dos variables aleatorias X, Y, está definida por

$$
corr(X, Y) = \frac{cov(X, Y)}{\sqrt{Var(X)Var(Y)}}.
$$

En este caso, al ser $Z^{(1)}$ y $Z^{(2)}$ $(\mathbb{G}, \mathbb{P})$-movimientos brownianos, sabemos que $\sqrt{Var(Z_t^{(i)})} = \sqrt{t}$, para $i = 1, 2$. Además, como $\mathbb{E}(Z_t^{(i)}) = 0$, para $i = 1, 2$, se tiene que

$$
\begin{aligned}
cov(Z_t^{(1)}, Z_t^{(2)}) &= \mathbb{E}(Z_t^{(1)} Z_t^{(2)})\\
&= \frac{1}{\sigma_1\sigma_2}\mathbb{E}\left(\sigma^2 (B_t^{(0)})^2 + \sigma B_t^{(0)}(b_1 B_t^{(1)} + b_2 B_t^{(2)}) + b_1 b_2 B_t^{(1)} B_t^{(2)}\right)\\
&= \frac{\sigma^2 t}{\sigma_1\sigma_2},
\end{aligned}
$$

donde se utilizó que $(B^{(0)}, B^{(1)}, B^{(2)})$ es un movimiento browniano tridimensional.

De lo anterior concluimos que

$$corr(Z_t^{(1)}, Z_t^{(2)}) = \frac{\sigma^2}{\sigma_1\sigma_2}.$$

■

Lema 4.1.2. *El proceso $\tilde{Z}^{(2)}$, definido por*

$$\tilde{Z}^{(2)} := \frac{Z^{(2)} - \rho Z^{(1)}}{\sqrt{1-\rho^2}},$$

es un $(\mathbb{G}, \mathbb{P})$*-movimiento browniano estándar.*

Demostración. Demostraremos que $\{(\tilde{Z}_t^{(2)}, \mathcal{G}_t), t \in [0,t]\}$ es martingala continua con variación cuadrática t. Luego, por el teorema de caracterización de Lévy 1.1.5 se sigue que $\tilde{Z}^{(2)}$ es $(\mathbb{G}, \mathbb{P})$-movimiento browniano.

Notemos que los procesos $Z^{(1)}$ y $Z^{(2)}$ son combinación lineal de movimientos brownianos y por ende son martingalas continuas respecto a la filtración $\mathbb{G}$, con $Z_0^{(1)} = 0$ y $Z_0^{(2)} = 0$. Por lo tanto, $\{(\tilde{Z}_t^{(2)}, \mathcal{G}_t), t \in [0,T]\}$ es una martingala continua, con $\tilde{Z}_0^{(2)} = 0$.

Por otro lado,

$$[\tilde{Z}^{(2)}]_t = \frac{1}{1-\rho^2}\Big([Z^{(2)}]_t - 2\rho[Z^{(1)}, Z^{(2)}]_t + \rho^2[Z^{(1)}]_t\Big) = \frac{1}{1-\rho^2}\Big(t - 2\rho^2 t + \rho^2 t\Big) = t,$$

con lo que concluimos el resultado. ■

Lema 4.1.3. *Los procesos $Z^{(1)}$ y $\tilde{Z}^{(2)}$ son independientes.*

Demostración. Ya que $Z^{(1)}$ y $\tilde{Z}^{(2)}$ son martingalas continuas con $Z_0^{(1)} = 0$ y $\tilde{Z}_0^{(2)} = 0$, y se satisface que $[\tilde{Z}^{(2)}]_t = t$, $[Z^{(1)}]_t = t$ y

$$\begin{aligned}
[Z^{(1)}, \tilde{Z}^{(2)}]_t &= \frac{1}{2}([Z^{(1)} + \tilde{Z}^{(2)}]_t - [Z^{(1)}]_t - [\tilde{Z}^{(2)}]_t) \\
&= \frac{1}{2}\left(\frac{t[1 + (\rho - \sqrt{1-\rho^2})^2 - 2\rho(\rho - \sqrt{1-\rho^2})]}{1-\rho^2} - t - t\right) \\
&= \frac{1}{2}\left(\frac{2t(1-\rho^2)}{1-\rho^2} - 2t\right) = 0,
\end{aligned}$$

por el Teorema 1.1.5 se sigue que $(Z^{(1)}, \tilde{Z}^{(2)})$ es $(\mathbb{G}, \mathbb{P})$-movimiento browniano bidimensional y en particular que $Z^{(1)}$ y $\tilde{Z}^{(2)}$ son independientes. ■

Definición 4.1.3. ***Proceso de innovación.*** *Definimos al proceso de innovación* $I = (I^{(1)}, I^{(2)})^T$, *para $t \in [0,T]$, como sigue*

$$I_t^{(1)} := Z_t^{(1)} + \int_0^t \frac{\mu_1(X_u,Y_u) - \boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du,$$
$$I_t^{(2)} := \tilde{Z}_t^{(2)} + \int_0^t \frac{\sigma_1[\mu_2(X_u,Y_u) - \boldsymbol{\mu}_2(X_u)^T\boldsymbol{\pi}_u] - \rho\sigma_2[\mu_1(X_u,Y_u) - \boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u]}{\sigma_1\sigma_2\sqrt{1-\rho^2}}du,$$

donde $\boldsymbol{\mu}_i(X_t)^T\boldsymbol{\pi}_t = \mathbb{E}(\mu_i(X_t,Y_t)|\mathcal{F}_t)$.

En forma matricial, para $t \in [0,T]$,

$$I_t = Z_t + \int_0^t \Sigma^{-1}\big[A(X_u,Y_u) - \pi_u(A(X_u,Y_u))\big]du,$$

con $Z_t = (Z^{(1)}, \tilde{Z}^{(2)})^T$, $A(X,Y) = (\mu_1(X,Y), \mu_2(X,Y))^T$ y

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma^2\sqrt{1-\rho^2} \end{pmatrix}.$$

Observación 4.1.2. *De la Definición 4.1.2, tenemos que*

$$I_t^{(1)} = \frac{R_t^{(1)}}{\sigma_1} - \int_0^t \frac{\boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du, \tag{4.5}$$

$$I_t^{(2)} = \frac{R_t^{(2)}}{\sigma_2\sqrt{1-\rho^2}} - \int_0^t \frac{\boldsymbol{\mu}_2(X_u)^T\boldsymbol{\pi}_u}{\sigma_2\sqrt{1-\rho^2}}du - \frac{\rho R_t^{(1)}}{\sigma_1\sqrt{1-\rho^2}} + \int_0^t \frac{\rho\boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1\sqrt{1-\rho^2}}du. \tag{4.6}$$

De lo anterior, heurísticamente, para δ pequeña, la observación $R_{t+\delta} - R_t$ tiene una parte que se puede predecir por $\delta\pi_s(A(X_s,Y_s))$ y un componente adicional $I_{t+\delta} - I_t$. Es por ello que al proceso estocástico I se le conoce como proceso de innovación, pues aporta información nueva a la obtenida mediante π.

En lo siguiente veremos que I es $(\mathbb{F},\mathbb{P})$-movimiento browniano y con ello obtendremos una representación de martingalas respecto a tal proceso. Tal resultado será de utilidad para probar la forma que tiene el proceso π.

Teorema 4.1.1. *El proceso $I = (I^{(1)}, I^{(2)})$ es $(\mathbb{F},\mathbb{P})$-movimiento browniano.*

Demostración. Utilizaremos el teorema de caracterización de Lévy, es decir, veremos que el proceso I es una martingala continua con $I_0 = 0$ y que se satisface que $[I^{(j)}]_t = t$, para $j = 1,2$ y $[I^{(1)}, I^{(2)}]_t = 0$.

1. I es martingala continua.

   Por definición se satisface que $I^{(1)}$ e $I^{(2)}$ son integrables. Además por (4.5) y (4.6) y la Observación 4.1.1 son también adaptados a $\mathbb{F}$.

Veamos ahora que se cumple la propiedad de martingala. Para $s \leq t$,

$$\begin{aligned}
\mathbb{E}(I_t^{(1)}|\mathcal{F}_s) - I_s^{(1)} &= \mathbb{E}\Big(Z_t^{(1)} + \int_0^t \frac{\mu_1(X_u,Y_u)}{\sigma_1}du\Big|\mathcal{F}_s\Big) - \mathbb{E}\Big(\int_0^t \frac{\boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du\Big|\mathcal{F}_s\Big) \\
&\quad - Z_s^{(1)} - \int_0^s \frac{\mu_1(X_u,Y_u)}{\sigma_1}du + \int_0^s \frac{\boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du \\
&= \mathbb{E}\Big(\frac{R_s^{(1)}}{\sigma_1} + Z_t^{(1)} - Z_s^{(1)} + \int_s^t \frac{\mu_1(X_u,Y_u)}{\sigma_1}du\Big|\mathcal{F}_s\Big) \\
&\quad - \mathbb{E}\Big(\int_0^t \frac{\boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du\Big|\mathcal{F}_s\Big) - \frac{R_s^{(1)}}{\sigma_1} + \int_0^s \frac{\boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du,
\end{aligned}$$

y al ser el proceso $R_s^{(1)}$ y $\int_0^s \boldsymbol{\mu_1}(X_u)^T\boldsymbol{\pi}_u du$ $\mathcal{F}_s$-medibles, obtenemos

$$\begin{aligned}
\mathbb{E}(I_t^{(1)}|\mathcal{F}_s) - I_s^{(1)} &= \mathbb{E}(Z_t^{(1)} - Z_s^{(1)}|\mathcal{F}_s) + \mathbb{E}\Big(\int_s^t \frac{\mu_1(X_u,Y_u)}{\sigma_1}du\Big|\mathcal{F}_s\Big) \\
&\quad - \mathbb{E}\Big(\int_s^t \frac{\boldsymbol{\mu}_1(X_u)^t\boldsymbol{\pi}_u}{\sigma_1}du\Big|\mathcal{F}_s\Big) \\
&= \mathbb{E}(\mathbb{E}(Z_t^{(1)} - Z_s^{(1)}|\mathcal{G}_s)|\mathcal{F}_s) + \mathbb{E}\Big(\int_s^t \frac{\mu_1(X_u,Y_u) - \boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du\Big|\mathcal{F}_s\Big) \\
&= \mathbb{E}\Big(\int_s^t \frac{\mu_1(X_u,Y_u) - \boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du\Big|\mathcal{F}_s\Big),
\end{aligned}$$

en donde se utilizó que $\mathbb{F}$ es sub σ-álgebra de $\mathbb{G}$ y que $Z^{(1)}$ es $(\mathbb{G},\mathbb{P})$-movimiento browniano.

Por otro lado, utilizando el teorema de Fubini

$$\begin{aligned}
&\mathbb{E}\Big(\int_s^t \frac{\mu_1(X_u,Y_u) - \boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}du\Big|\mathcal{F}_s\Big) \\
=&\int_s^t \mathbb{E}\Big(\frac{\mu_1(X_u,Y_u) - \boldsymbol{\mu}_1(X_u)^T\boldsymbol{\pi}_u}{\sigma_1}\Big|\mathcal{F}_s\Big)du \\
=&\int_s^t \frac{1}{\sigma_1}\Big[\mathbb{E}(\mu_1(X_u,Y_u)|\mathcal{F}_s) - \mathbb{E}(\mathbb{E}(\mu_1(X_u,Y_u)|\mathcal{F}_u)|\mathcal{F}_s)\Big]du \\
=&0,
\end{aligned}$$

de donde se sigue que $I^{(1)}$ es martingala continua.

Para $I^{(2)}$ notemos que

$$I_t^{(2)} = \frac{R_t^{(2)}}{\sigma_2\sqrt{1-\rho^2}} - \int_0^t \frac{\boldsymbol{\mu}_2(X_u)^T\boldsymbol{\pi}_u}{\sigma_2\sqrt{1-\rho^2}}du - \frac{\rho I_t^{(1)}}{\sqrt{1-\rho^2}},$$

por lo tanto, análogamente al caso anterior podemos ver que

$$\frac{R_t^{(2)}}{\sigma_2\sqrt{1-\rho^2}} - \int_0^t \frac{\boldsymbol{\mu}_2(X_u)^T\boldsymbol{\pi}_u}{\sigma_2\sqrt{1-\rho^2}}du$$

es $(\mathbb{F},\mathbb{P})$-martingala y así $I^{(2)}$ también lo es.

Con ello concluimos que I es martingala continua y es claro que $I_0 = 0$.

2. Como $Z^{(1)}$ y $\tilde{Z}^{(2)}$ son $(\mathbb{G},\mathbb{P})$-movimientos brownianos independientes, por definición del proceso I, se tiene que $[I^{(1)}]_t = t$, $[I^{(2)}]_t = t$ y $[I^{(1)}, I^{(2)}]_t = 0$.

De lo anterior se sigue el resultado. ■

Observación 4.1.3. *De las ecuaciones* (4.5) *y* (4.6) *podemos concluir que* $(S^{(m)}, R^{(1)}, R^{(2)})$ *y* $(S^{(m)}, I^{(1)}, I^{(2)})$ *generan la misma* σ*-álgebra, la prueba se puede consultar en [2].*

Teorema 4.1.2. *Supongamos que* I *es un proceso de Wiener en* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$*. Entonces toda* $(\mathbb{F},\mathbb{P})$*-martingala local tiene la siguiente representación* $M = M_0 + H \cdot I$*, para algún proceso* H *localmente cuadrado integrable.*

Demostración. Ver Jacod y Shiryaev [6, Teorema 3.4.34]. ■

Observación 4.1.4. *Vimos que el proceso* I *es* $(\mathbb{F},\mathbb{P})$*-movimiento browniano entonces, por el Teorema 4.1.2 y la Observación 4.1.3, toda* $(\mathbb{F},\mathbb{P})$*-martingala local* M *tiene la siguiente representación*

$$M_t = M_0 + \int_0^t \gamma_u dI_u,$$

con $t \in [0,T]$ *y para algún proceso* γ *bidimensional y* $\mathbb{F}$*-predecible tal que*

$$\int_0^T ||\gamma_u||^2 du < \infty,$$

$\mathbb{P}$*-c.s.*

Con los resultados obtenidos podemos ahora demostrar el resultado de interés.

Proposición 4.1.1. *Para cada* $i \in \mathcal{E}$*, el proceso* π_t^i *satisface el siguiente sistema de ecuaciones diferenciales estocásticas*

$$d\pi_t^i = \sum_{j\in\mathcal{E}} q^{ij}\pi_t^j dt + H^i(X_t, \boldsymbol{\pi}_t) dI_t,$$

con $\pi_0^i = p_0$ *y para* $i \in \mathcal{E}$*, el proceso bidimensional* $H^i(X, \boldsymbol{\pi})$ *está dado por*

$$\begin{aligned} H^{i,(1)}(X_t, \boldsymbol{\pi}_t) &= \frac{\pi_t^i(\mu_1(X_t, e_i) - \boldsymbol{\mu}_1(X_t)^T \boldsymbol{\pi}_t)}{\sigma_1}, \\ H^{i,(2)}(X_t, \boldsymbol{\pi}_t) &= \frac{\pi_t^i[\sigma_1(\mu_2(X_t, e_i) - \boldsymbol{\mu}_2(X_t)^T\boldsymbol{\pi}_t) - \sigma_2\rho(\mu_1(X_t, e_i) - \boldsymbol{\mu}_1(X_t)^T\boldsymbol{\pi}_t)]}{\sigma_1\sigma_2\sqrt{1-\rho^2}}, \end{aligned}$$

para $t \in [0, T]$ *con* $\sigma_1 = \sqrt{\sigma^2 + b_1^2}$ *y* $\sigma_2 = \sqrt{\sigma_2 + b_2^2}$.

Para la prueba de la proposición anterior utilizaremos el siguiente resultado:

Proposición 4.1.2. *Sea* Y *una cadena de Markov continua con espacio de estados* $\mathcal{E} = \{e_1, \ldots, e_k\}$. *Para toda función* $f : \mathcal{E} \to \mathbb{R}$ *Borel medible y acotada, el proceso* $M^f = \{M_t^f, t \geq 0\}$ *definido por*

$$M_t^f = f(Y_t) - f(Y_0) - \int_0^t Qf(Y_s)ds,$$

es una $(\mathbb{G}, \mathbb{P})$*-martingala, donde* Q *el generador de la cadena* Y*, dado por* $Qf(i) = \sum_{j \in \mathcal{E}} q^{ij} f(j)$.

Demostración. Notemos que si definimos, para $j \in \mathcal{E}$, al proceso $M^j = \{M_t^j, t \geq 0\}$ como

$$M_t^j = \mathbb{1}_{\{Y_t = j\}} - \mathbb{1}_{\{Y_0 = j\}} - \int_0^t q^{Y_s j} ds,$$

entonces podemos escribir a M^f como

$$M^f = \sum_{j \in \mathcal{E}} f(j) M^j.$$

Así, basta probar que el proceso M^j, para $j \in \mathcal{E}$, es $(\mathbb{G}, \mathbb{P})$-martingala.

Por definición de M_t^j, es claro que es adaptado a $\mathbb{G}$ y además M_t^j es integrable. Veamos que se cumple la propiedad de martingala, para $s \leq t$, tenemos lo siguiente

$$\mathbb{E}(M_t^j | \mathcal{G}_s) = \mathbb{P}(Y_t = j | \mathcal{G}_s) - \mathbb{1}_{\{Y_0 = j\}} - \mathbb{E}\left(\int_0^t q^{Y_r j} dr \middle| \mathcal{G}_s \right).$$

Por otro lado, recordemos que por la ecuación de Kolmogorov hacía adelante se satisface que

$$\mathbb{P}(Y_t = j | Y_s = i) = \delta_{ij} + \int_s^t \sum_{l \in \mathcal{E}} q^{lj} \mathbb{P}(Y_r = l | Y_s = i) dr,$$

así, utilizando la expresión anterior y la por la propiedad de Markov se sigue que

$$\begin{aligned}
\mathbb{P}(Y_t = j | \mathcal{G}_s) &= \mathbb{P}(Y_t = j | Y_s) \\
&= \sum_{i \in \mathcal{E}} \mathbb{P}(Y_t = j | Y_s = i) \mathbb{1}_{\{Y_s = i\}} \\
&= \sum_{i \in \mathcal{E}} \mathbb{1}_{\{Y_s = i\}} \left(\delta_{ij} + \int_s^t \sum_{l \in \mathcal{E}} q^{lj} \mathbb{P}(Y_r = l | Y_s = i) dr \right) \\
&= \mathbb{1}_{\{Y_s = j\}} + \int_s^t \sum_{i \in \mathcal{E}} \sum_{l \in \mathcal{E}} \mathbb{1}_{\{Y_s = i\}} q^{lj} \mathbb{P}(Y_r = l | Y_s = i) dr
\end{aligned}$$

$$= \mathbb{1}_{\{Y_s=j\}} + \mathbb{E}\left(\int_s^t q^{Y_r j} dr \Big| \mathcal{G}_s \right).$$

Por lo tanto,

$$\begin{aligned}
\mathbb{E}(M_t^j | \mathcal{G}_s) &= \mathbb{1}_{\{Y_s=j\}} + \mathbb{E}\left(\int_s^t q^{Y_r j} dr \Big| \mathcal{G}_s \right) - \mathbb{1}_{\{Y_0=j\}} - \mathbb{E}\left(\int_0^t q^{Y_r j} dr \Big| \mathcal{G}_s \right) \\
&= \mathbb{1}_{\{Y_s=j\}} - \mathbb{1}_{\{Y_0=j\}} - \mathbb{E}\left(\int_0^s q^{Y_r j} dr \Big| \mathcal{G}_s \right) \\
&= M_s^j.
\end{aligned}$$

Con lo que concluimos que, para $j \in \mathcal{E}$, M^j es $(\mathbb{G}, \mathbb{P})$-martingala y así el proceso M^f también lo es. ■

Demostración. (De la Proposición 4.1.1).
Por la Proposición 4.1.2, para $f : \mathcal{E} \to \mathbb{R}$ Borel medible y acotada, tenemos que el proceso M_t^1, definido por

$$M_t^1 = f(Y_t) - f(Y_0) - \int_0^t Qf(Y_s) ds, \quad t \in [0, T], \tag{4.7}$$

es $(\mathbb{G}, \mathbb{P})$-martingala.

Entonces,

$$\mathbb{E}(M_t^1 | \mathcal{F}_t) = \mathbb{E}(f(Y_t) | \mathcal{F}_t) - \mathbb{E}(f(Y_0) | \mathcal{F}_t) - \mathbb{E}\left(\int_0^t Qf(Y_s) ds \Big| \mathcal{F}_t \right), \quad t \in [0, T].$$

Notemos que $M_t^2 = \mathbb{E}(M_t^1 | \mathcal{F}_t)$ es $(\mathbb{F}, \mathbb{P})$-martingala, pues es integrable ya que M^1 lo es y por definición es adaptada a $\mathbb{F}$. Además, para $s \leq t$

$$\begin{aligned}
\mathbb{E}\big[M_t^2 | \mathcal{F}_s \big] &= \mathbb{E}\big[\mathbb{E}(M_t^1 | \mathcal{F}_t) | \mathcal{F}_s \big] \\
&= \mathbb{E}(M_t^1 | \mathcal{F}_s) \\
&= \mathbb{E}\big[\mathbb{E}(M_t^1 | \mathcal{G}_s) | \mathcal{F}_s \big] \\
&= \mathbb{E}(M_s^1 | \mathcal{F}_s) \\
&= M_s^2,
\end{aligned}$$

donde se utilizó que $\mathbb{F}$ es sub σ-álgebra de $\mathbb{G}$.

Por lo tanto, por la Observación 4.1.4, tenemos lo siguiente

$$\mathbb{E}\big[f(Y_t) | \mathcal{F}_t \big] - \mathbb{E}\big[f(Y_0) | \mathcal{F}_t \big] - \mathbb{E}\left(\int_0^t Qf(Y_s) ds \Big| \mathcal{F}_t \right) = \int_0^t \gamma_u dI_u, \quad t \in [0, T], \tag{4.8}$$

para algún proceso γ bidimensional y $\mathbb{F}$-predecible tal que

$$\int_0^T ||\gamma_u||^2 du < \infty, \quad \mathbb{P} - \text{c.s.}$$

Consideremos ahora el proceso

$$m_t = I_t + \int_0^t \Sigma^{-1} \pi_u(A(X_u, Y_u)) du, \quad t \in [0, T], \tag{4.9}$$

donde I_t, el proceso de innovación y $A(X,Y)$ están definidos en (4.1.3).

Notemos que el proceso m es adaptado a la filtración $\mathbb{F}$. Además, por la Definición 4.1.3 lo podemos escribir como

$$m_t = Z_t + \int_0^t \Sigma^{-1} A(X_u, Y_u) du. \tag{4.10}$$

Luego, de las ecuaciones de (4.7) y (4.10), aplicando la fórmula de integración por partes a los componentes $f(Y_t)m_t = (f(Y_t)m_t^{(1)}, f(Y_t)m_t^{(2)})$ llegamos a

$$d(f(Y_t)m_t^{(1)}) = m_t^{(1)} Qf(Y_t)dt + m_t^{(1)} dM_t^1 + f(Y_t)dZ_t^{(1)} + f(Y_t)(\Sigma^{-1}A(X_t, Y_t))_1 dt + d[Z^{(1)}, M^1]_t,$$
$$d(f(Y_t)m_t^{(2)}) = m_t^{(2)} Qf(Y_t)dt + m_t^{(2)} dM_t^1 + f(Y_t)d\tilde{Z}_t^{(2)} + f(Y_t)(\Sigma^{-1}A(X_t, Y_t))_2 dt + d[\tilde{Z}^{(2)}, M^1]_t,$$

donde $(\Sigma^{-1}A(X_t, Y_t))_k$ denota la entrada k del vector $\Sigma^{-1}A(X_t, Y_t)$, para $k = 1, 2$. Por lo tanto,

$$d(f(Y_t)m_t) = m_t Qf(Y_t)dt + m_t dM_t^1 + f(Y_t)dZ_t + \Sigma^{-1} f(Y_t)A(X_t, Y_t)dt + d[Z, M^1]_t,$$

esto es,

$$f(Y_t)m_t = \int_0^t m_s Qf(Y_s)ds + \int_0^t m_s dM_s^1 + \int_0^t f(Y_s)dZ_s + \int_0^t \Sigma^{-1} f(Y_s)A(X_s, Y_s)ds + [Z, M^1]_t.$$

Proyectando sobre la filtración $\mathcal{F}_t$, obtenemos

$$\mathbb{E}\big[f(Y_t)m_t | \mathcal{F}_t\big] = \mathbb{E}\bigg(\int_0^t m_s Qf(Y_s)ds \big| \mathcal{F}_t\bigg) + \mathbb{E}\bigg(\int_0^t \Sigma^{-1} f(Y_s)A(X_s, Y_s)ds \big| \mathcal{F}_t\bigg) + M_t^3, \tag{4.11}$$

donde, M_t^3 es $(\mathbb{F}, \mathbb{P})$ martingala, definida por

$$M_t^3 = \mathbb{E}\bigg(\int_0^t m_s dM_s^1 \big| \mathcal{F}_t\bigg) + \mathbb{E}\bigg(\int_0^t f(Y_s)dZ_s \big| \mathcal{F}_t\bigg) + \mathbb{E}([Z, M^1]_t | \mathcal{F}_t).$$

Por otro lado, de la ecuación (4.8), por el teorema de Fubini se tiene que

$$\pi_t(f) = \mathbb{E}(f(Y_t)|\mathcal{F}_t) = \mathbb{E}(f(Y_0)|\mathcal{F}_t) + \int_0^t \mathbb{E}(Qf(Y_s)|\mathcal{F}_t)ds + \int_0^t \gamma_u dI_u, \quad t \in [0, T]. \tag{4.12}$$

Análogamente, aplicado integración por partes a los componentes del producto $\pi_t(f)m_t$, de (4.9) y (4.12), obtenemos que

$$d(\pi_t(f)m_t) = m_t\mathbb{E}(Qf(Y_t)|\mathcal{F}_t)dt + m_t\gamma_t dI_t + \pi_t(f)dI_t + \Sigma^{-1}\pi_t(f)\pi_t(A(X_t,Y_t))dt + \gamma_t dt,$$

de forma integral,

$$\begin{aligned}\pi_t(f)m_t &= \mathbb{E}(f(Y_t)|\mathcal{F}_t)m_t \\ &= \int_0^t m_s\mathbb{E}(Qf(Y_s)|\mathcal{F}_s)ds + \int_0^t \Sigma^{-1}\pi_s(f)\pi_s(A(X_s,Y_s))ds + \int_0^t \gamma_s ds + M_t^4, \qquad (4.13)\end{aligned}$$

donde M_t^4 es $(\mathbb{F},\mathbb{P})$ martingala local, y está dada por

$$M_t^4 = \int_0^t m_s\gamma_s dI_s + \int_0^t \pi_s(f)dI_s.$$

Luego, como $\mathbb{E}(f(Y_t)m_t|\mathcal{F}_t) = m_t\mathbb{E}(f(Y_t)|\mathcal{F}_t)$, por ser m adaptado a $\mathbb{F}$, comparando los términos de variación finita de (4.11) y (4.13), obtenemos que, para $t \in [0,T]$

$$\gamma_t = \Sigma^{-1}\big[\mathbb{E}(f(Y_t)A(X_t,Y_t)|\mathcal{F}_t) - \mathbb{E}(f(Y_t)|\mathcal{F}_t)\mathbb{E}(A(X_t,Y_t)|\mathcal{F}_t)\big],$$

esto es,

$$\begin{aligned}\gamma_t^{(1)} &= \frac{\mathbb{E}(f(Y_t)\mu_1(X_t,Y_t)|\mathcal{F}_t) - \mathbb{E}(f(Y_t)|\mathcal{F}_t)\mathbb{E}(\mu_1(X_t,Y_t)|\mathcal{F}_t)}{\sigma_1}, \\ \gamma_t^{(2)} &= \frac{\sigma_1[\mathbb{E}(f(Y_t)\mu_2(X_t,Y_t)|\mathcal{F}_t) - \mathbb{E}(f(Y_t)|\mathcal{F}_t)\mathbb{E}(\mu_2(X_t,Y_t)|\mathcal{F}_t)]}{\sigma_1\sigma_2\sqrt{1-\rho^2}} \\ &\quad - \frac{\sigma_2\rho[\mathbb{E}(f(Y_t)\mu_1(X_t,Y_t)|\mathcal{F}_t) - \mathbb{E}(f(Y_t)|\mathcal{F}_t)\mathbb{E}(\mu_1(X_t,Y_t)|\mathcal{F}_t)]}{\sigma_1\sigma_2\sqrt{1-\rho^2}},\end{aligned}$$

ya que

$$\Sigma^{-1} = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}\begin{pmatrix}\sigma_2\sqrt{1-\rho^2} & 0 \\ -\sigma_2\rho & \sigma_1\end{pmatrix}.$$

Ahora, si consideramos $f(Y_t) = \mathbb{1}_{\{Y_t=e_i\}}$, se tiene que

$$\begin{aligned}\gamma_t^{(1)} &= \frac{\pi_t^i\mu_1(X_t,e_i) - \pi_t^i\mathbb{E}(\mu_1(X_t,Y_t)|\mathcal{F}_t)}{\sigma_1}, \\ \gamma_t^{(2)} &= \frac{\sigma_1[\pi_t^i\mu_2(X_t,e_i) - \pi_t^i\mathbb{E}(\mu_2(X_t,Y_t)|\mathcal{F}_t)] - \sigma_2\rho[\pi_t^i\mu_1(X_t,e_i) - \pi_t^i\mathbb{E}(\mu_1(X_t,Y_t)|\mathcal{F}_t)]}{\sigma_1\sigma_2\sqrt{1-\rho^2}},\end{aligned}$$

pues $\mathbb{E}(\mathbb{1}_{\{Y_t=e_i\}}\mu_k(X_t,Y_t)|\mathcal{F}_t) = \mathbb{P}(Y_t=e_i|\mathcal{F}_t)\mu_k(X_t,e_i) = \pi_t^i\mu_k(X_t,e_i)$, para $k=1,2$.

Así, de la ecuación (4.8), la dinámica de π_t^i, para cada $i \in \mathcal{E}$, está dada por

$$d\pi_t^i = \sum_{j \in \mathcal{E}} q^{ij} \pi_t^j dt + H^i(X_t, \boldsymbol{\pi}_t) dI_t,$$

con $H^i(X_t, \boldsymbol{\pi_t})$ como en el enunciado del teorema.

Además, al ser los términos de deriva y de difusión continuos, acotados y localmente Lipschitz, se tiene existencia y unicidad de la solución del sistema de ecuaciones diferenciales estocásticas anterior. ■

Finalmente, podemos ahora reescrbir la dinámica de los activos cointegrados en términos de la dinámica del proceso $\boldsymbol{\pi}$. Observemos primero que, de las ecuaciones (4.5) y (4.6) y por la Definición 4.1.2 se tiene que

$$\begin{aligned}
I_t^{(1)} &= \frac{\sigma B_t^{(0)} + b_1 B_t^{(1)}}{\sigma_1} - \int_0^t \frac{\boldsymbol{\mu}_1(X_u)^T \boldsymbol{\pi}_u}{\sigma_1} du, \\
I_t^{(2)} &= \frac{\sigma B_t^{(0)} + b_2 B_t^{(2)}}{\sigma_2 \sqrt{1-\rho^2}} - \int_0^t \frac{\boldsymbol{\mu}_2(X_u)^T \boldsymbol{\pi}_u}{\sigma_2 \sqrt{1-\rho^2}} du - \frac{\rho[\sigma B_t^{(0)} + b_1 B_t^{(1)}]}{\sigma_1 \sqrt{1-\rho^2}} + \int_0^t \frac{\rho \boldsymbol{\mu}_1(X_u)^T \boldsymbol{\pi}_u}{\sigma_1 \sqrt{1-\rho^2}} du.
\end{aligned}$$

Por lo anterior y las ecuaciones (2.2) y (2.3) obtenemos que

$$\frac{dS_t^{(1)}}{S_t^{(1)}} = (r + \beta_1 \mu_m) dt + \beta_1 \sigma_m dB_t^{(m)} + \sigma_1 dI_t^{(1)} + \boldsymbol{\mu}_1(X_t)^T \boldsymbol{\pi}_t dt, \tag{4.14}$$

$$\frac{dS_t^{(2)}}{S_t^{(2)}} = (r + \beta_2 \mu_m) dt + \beta_2 \sigma_m dB_t^{(m)} + \sigma_2 \rho dI_t^{(1)} + \sigma_2 \sqrt{1-\rho^2} dI_t^{(2)} + \boldsymbol{\mu}_2(X_t)^T \boldsymbol{\pi}_t dt, \tag{4.15}$$

con $S_0^{(1)} > 0$ y $S_0^{(2)} > 0$. Notemos que la dinámica de $S^{(m)}$ no cambia pues no está afectada por la cadena Y.

Así, aplicando la fórmula de Itô al proceso $X_t = \log S_t^{(1)} - \log S_t^{(2)}$, llegamos a

$$\begin{aligned}
dX_t &= \mu_m(\beta_1 - \beta_2) dt + \sigma_m(\beta_1 - \beta_2) dB_t^{(m)} + \sigma_1 dI_t^{(1)} - \sigma_2 \rho dI_t^{(1)} + \boldsymbol{\mu}_1(X_t)^T \boldsymbol{\pi}_t \\
&\quad - \sigma_2 \sqrt{1-\rho^2} dI_t^{(2)} - \boldsymbol{\mu}_2(X_t)^T \boldsymbol{\pi}_t dt - \frac{1}{2}\left[\beta_1^2 \sigma_m^2 + \sigma_1^2 - \beta_2^2 \sigma_m^2 - \sigma_2^2 \rho^2 - \sigma_2^2(1-\rho^2)\right] dt \\
&= \left[\Gamma_1 + \boldsymbol{\mu}_1(X_t)^T \boldsymbol{\pi}_t - \boldsymbol{\mu}_2(X_t)^T \boldsymbol{\pi}_t\right] dt + (\beta_1 - \beta_2) \sigma_m dB_t^{(m)} \\
&\quad + (\sigma_1 - \rho \sigma_2) dI_t^{(1)} - \sigma_2 \sqrt{1-\rho^2} dI_t^{(2)},
\end{aligned} \tag{4.16}$$

con $\Gamma_1 = (\beta_1 - \beta_2)\sigma_m - \frac{1}{2}\left[(\beta_1 - \beta_2)\sigma_m^2 + b_1^2 - b_2^2\right]$.

Para las siguientes secciones se supone que $\lambda_k(y) = \lambda_k$, para $k = 1, 2$. Esto permitirá poder obtener una forma explícita de la función de valor módulo la solución de un sistema de ecuaciones diferenciales parciales. Para este caso los coeficientes de difusión no dependerán del proceso X, esto es, para $t \in [0, T]$,

$$H^{i,(1)}(X_t, \boldsymbol{\pi}_t) = H^{i,(1)}(\boldsymbol{\pi}_t) = \frac{\lambda_1 \pi_t^i [\alpha_1^i - \boldsymbol{\alpha}_1^T \boldsymbol{\pi}_t]}{\sigma_1},$$
$$H^{i,(2)}(X_t, \boldsymbol{\pi}_t) = H^{i,(2)}(\boldsymbol{\pi}_t) = \frac{-\lambda_2 \sigma_1 \pi_t^i (\alpha_2^i - \boldsymbol{\alpha}_2^T \boldsymbol{\pi}_t) - \sigma_2 \lambda_1 \rho \pi_t^i (\alpha_1^i - \boldsymbol{\alpha}_2^T \boldsymbol{\pi}_t)}{\sigma_1 \sigma_2 \sqrt{1-\rho^2}},$$

ya que, al no haber dependencia del proceso X, para cada $i \in \mathcal{E}$, $\mu_1(X_t, e_i) = \lambda_1 \alpha_1(e_i) =: \lambda_1 \alpha_1^i$ y $\mu_2(X_t, e_i) = -\lambda_2 \alpha_2(e_i) =: -\lambda_2 \alpha_2^i$. Y además se satisface que

$$\boldsymbol{\mu}_1(X_t)^T \boldsymbol{\pi}_t = \lambda_1 \mathbb{E}(\alpha_1(Y_t)|\mathcal{F}_t) = \lambda_1(\alpha_1(e_1), \ldots, \alpha_1(e_k))\boldsymbol{\pi}_t =: \lambda_1 \boldsymbol{\alpha}_1^T \boldsymbol{\pi}_t,$$
$$\boldsymbol{\mu}_2(X_t)^T \boldsymbol{\pi}_t = -\lambda_2 \mathbb{E}(\alpha_2(Y_t)|\mathcal{F}_t) = -\lambda_2(\alpha_2(e_1), \ldots, \alpha_2(e_k))\boldsymbol{\pi}_t =: -\lambda_2 \boldsymbol{\alpha}_2^T \boldsymbol{\pi}_t.$$

4.2. Caso clásico

Para el caso clásico se supone que no hay restricción sobre las carteras de inversión. Para un agente que desee invertir su riqueza en un mercado como el ya descrito definiremos a las estrategias admisibles como sigue:

Definición 4.2.1. *Una estrategia $h = (h^{(m)}, h^{(1)}, h^{(2)})$ es $\mathbb{F}$-admisible si es una estrategia autofinanciada y $\mathbb{F}$-predecible y que satisface la condición de integrabilidad* (3.1)*. Denotaremos por $\mathcal{A}^{\mathbb{F}}$ al conjunto de las estrategias $\mathbb{F}$-admisibles.*

Así, la dinámica del portafolio W^h, asociado a la estrategia $h = (h^{(m)}, h^{(1)}, h^{(2)})$, estará dado por los intereses generados más el capital ganado debido a cambios en los activos $S^{(m)}$, $S^{(1)}$ y $S^{(2)}$, es decir,

$$dW_t^h = W_t^h \left[h_t^{(1)} \frac{dS_t^{(1)}}{S_t^{(1)}} + h_t^{(2)} \frac{dS_t^{(2)}}{S_t^{(2)}} + h_t^{(m)} \frac{dS_t^{(m)}}{S_t^{(m)}} \right] + r W_t^h \big[1 - h_t^{(1)} - h_t^{(2)} - h_t^{(m)}\big] dt,$$

y por (2.1), (4.14) y (4.15), obtenemos

$$\begin{aligned} dW_t^h &= W_t^h \Big(\big[r + (h_t^{(m)} + h_t^{(1)} \beta_1 + h_t^{(2)} \beta_2) \mu_m + h_t^{(1)} \boldsymbol{\mu}_1(X_t)^T \boldsymbol{\pi}_t + h_t^{(2)} \boldsymbol{\mu}_2(X_t)^T \boldsymbol{\pi}_t \big] dt \\ &\qquad + \sigma_m (h_t^{(m)} + h_t^{(1)} \beta_1 + h_t^{(2)} \beta_2) dB_t^{(m)} + (\sigma_1 h_t^{(1)} + \rho \sigma_2 h_t^{(2)}) dI_t^{(1)} + \sigma_2 \sqrt{1-\rho^2} h_t^{(2)} dI_t^{(2)} \Big) \\ &= f(t, X_t, \boldsymbol{\pi}_t) W_t^h dt + \sum_{l=0}^{2} g_l(t) W_t^h dB_t^l, \end{aligned} \tag{4.17}$$

con $W_0^h > 0$ y donde $f(t, X_t, \boldsymbol{\pi}_t) := r + (h_t^{(m)} + h_t^{(1)} \beta_1 + h_t^{(2)} \beta_2) \mu_m + h_t^{(1)} \boldsymbol{\mu}_1(X_t)^T \boldsymbol{\pi}_t + h_t^{(2)} \boldsymbol{\mu}_2(X_t)^T \boldsymbol{\pi}_t$,

$(B^0, B^1, B^2) \coloneqq (B^{(m)}, I^{(1)}, I^{(2)})$ y

$$\begin{aligned}
g_0(t) &\coloneqq \sigma_m(h_t^{(m)} + h_t^{(1)}\beta_1 + h_t^{(2)}\beta_2),\\
g_1(t) &\coloneqq \sigma_1 h_t^{(1)} + \rho\sigma_2 h_t^{(2)},\\
g_2(t) &\coloneqq \sigma_2\sqrt{1-\rho^2}h_t^{(2)}.
\end{aligned}$$

Por lo que, usando el Lema 1.1.1, para $u \leq t$, podemos escribir a W^h de la siguiente forma

$$\begin{aligned}
W_t^h &= W_u^h \exp\left(\int_u^t \left[f(s, X_x, \boldsymbol{\pi}_s) - \frac{1}{2}\sum_{l=0}^{2} g_l^2(s)\right]ds + \sum_{l=0}^{2}\int_u^t g_l(s)dB_s^l\right)\\
&=: W_u^h \exp(G(u, t, X_s, \boldsymbol{\pi}_s; u \leq s \leq t)),
\end{aligned} \tag{4.18}$$

donde

$$G(u, t, X_s, \boldsymbol{\pi}_s; u \leq s \leq t) \coloneqq \int_u^t \left[f(s, X_x, \boldsymbol{\pi}_s) - \frac{1}{2}\sum_{l=0}^{2} g_l^2(s)\right]ds + \sum_{l=0}^{2}\int_u^t g_l(s)dB_s^l.$$

4.2.1. Problema de control

En la sección 4.1 se obtuvo la existencia y unicidad del proceso π y con esto se expresó la dinámica de los procesos X y W respecto a la filtración $\mathbb{F}$. Esto nos permite plantear, en términos del proceso de estados (W, X, π), el problema de control equivalente bajo información completa. Para ello, definiremos primero el siguiente operador:

Definición 4.2.2. *Sea $F(t, w, x, \boldsymbol{p})$ en $C^{1,2,2,2}([0,T], \mathbb{R}^+, \mathbb{R}, \Delta_k)$, definimos al operador de segundo orden*

$$\begin{aligned}
&\mathcal{L}_{\mathbb{F}}^h F(t, w, x, \boldsymbol{p})\\
\coloneqq &F_t(t, w, x, \boldsymbol{p}) + F_x(t, w, x, \boldsymbol{p})[\Gamma_1 + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}] + \sum_{i,j\in\mathcal{E}} F_{p_i}(t, w, x, \boldsymbol{p})q^{ij}p^j\\
&+ \left[r + (h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)\mu_m + h^{(1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} + h^{(2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p}\right]wF_w(t, w, x, \boldsymbol{p})\\
&+ \frac{1}{2}F_{xx}(t, w, x, \boldsymbol{p})\Gamma_2 + \frac{1}{2}\sum_{i,j\in\mathcal{E}} F_{p_ip_j}(t, w, x, \boldsymbol{p})\left[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\right]\\
&+ \frac{1}{2}F_{ww}(t, w, x, \boldsymbol{p})w^2\big[(h^{(m)} + \beta_1 h^{(1)} + \beta_2 h^{(2)})^2\sigma_m^2 + (\sigma_1 h^{(1)} + \rho\sigma_2 h^{(2)})^2\\
&\qquad\qquad + \sigma_2^2(1-\rho^2)(h^{(2)})^2\big]\\
&+ F_{wx}(t, w, x, \boldsymbol{p})w\big[\sigma_m^2(\beta_1 - \beta_2)(h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2) + \sigma_1^2 h^{(1)} - \sigma_2^2 h^{(2)}\\
&\qquad\qquad - \rho\sigma_1\sigma_2(h^{(1)} - h^{(2)})\big]\\
&+ \sum_{i\in\mathcal{E}} F_{wp_i}(t, w, x, \boldsymbol{p})w\left[H^{i,(1)}(\boldsymbol{p})(\sigma_1 h^{(1)} + \rho\sigma_2 h^{(2)}) + H^{i,(2)}(\boldsymbol{p})\sqrt{1-\rho^2}\sigma_2 h^{(2)}\right]
\end{aligned}$$

$$+\sum_{i\in\mathcal{E}} F_{xp_i}(t,w,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1-\rho\sigma_2)-H^{i,(2)}(\boldsymbol{p})\sigma_2\sqrt{1-\rho^2}\big], \tag{4.19}$$

con $\Gamma_2=\sigma_m^2(\beta_1-\beta_2)^2+b_1^2+b_2^2$, $\sigma_1=\sqrt{\sigma^2+b_1^2}$ *y* $\sigma_2=\sqrt{\sigma^2+b_2^2}$, *donde* $\Delta_k:=\{(p_1,\ldots,p_k)^T:\sum_{1\leq i\leq k}p_i=1,p_i\geq 0,\forall i\}$.

Consideremos entonces un inversionista que desea maximizar su utilidad esperada asignando su riqueza a un mercado como el descrito y además con función de utilidad logarítmica. Tenemos así el siguiente problema de control:

$$\text{maximizar}\quad \mathbb{E}_{t,w,x,\boldsymbol{p}}(\log W_t^h),\quad \text{sobre todas las estrategias } h\in\mathcal{A}^{\mathbb{F}},$$

donde $\mathbb{E}_{t,w,x,\boldsymbol{p}}$ denota la esperanza condicional dado que $(W_t,X_t,\boldsymbol{\pi}_t)=(w,x,\boldsymbol{p})$, con $w\in\mathbb{R}^+$, $x\in\mathbb{R}$ y $\boldsymbol{p}\in\Delta_k$.

En este contexto definimos a la función de valor como sigue

$$V(t,w,x,\boldsymbol{p}):=\sup_{h\in\mathcal{A}^{\mathbb{F}}}\mathbb{E}_{t,w,x,\boldsymbol{p}}(\log W_T^h).$$

Teorema 4.2.1. *Supongamos que* $\lambda_1(y)=\lambda_1\in\mathbb{R}$, $\lambda_2(y)=\lambda_2\in\mathbb{R}$ *y consideremos una inversión con función de utilidad logarítmica. Entonces la estrategia óptima* $h_*=(h_*^{(m)},h_*^{(1)},h_*^{(2)})\in\mathcal{A}^{\mathbb{F}}$ *es*

$$\begin{aligned}
h_*^{(1)}(t,x,\boldsymbol{p})&=\frac{\boldsymbol{\mu}_1(x)^T\boldsymbol{p}-\boldsymbol{\mu}_2(x)^T\boldsymbol{p}\varrho_2}{b_1^2+b_2^2\varrho_2},\\
h_*^{(2)}(t,x,\boldsymbol{p})&=\frac{\boldsymbol{\mu}_2(x)^T\boldsymbol{p}-\boldsymbol{\mu}_1(x)^T\boldsymbol{p}\varrho_1}{b_2^2+b_1^2\varrho_1},\\
h_*^{(m)}(t,x,\boldsymbol{p})&=\frac{\mu_m}{\sigma_m^2}-\beta_1h_*^{(1)}(t,w,\boldsymbol{p})-\beta_2h_*^{(2)}(t,w,\boldsymbol{p}).
\end{aligned}$$

Además, la función de valor es de la forma $V(t,w,x,\boldsymbol{p})=\log(w)+\bar{m}(t)x^2+\bar{n}(t,\boldsymbol{p})x+\bar{u}(t,\boldsymbol{p})$, *donde la función* $\bar{m}(t)$, *para* $0\leq t\leq T$, *resuelve la ecuación diferencial ordinaria*

$$\bar{m}_t(t)-2\bar{m}(t)(\lambda_1+\lambda_2)+\Theta_1=0, \tag{4.20}$$

con condición terminal $\bar{m}(T)=0$. *Y las funciones* $\bar{n}(t,\boldsymbol{p})$ *y* $\bar{u}(t,\boldsymbol{p})$, *para* $0\leq t\leq T$, *resuelven el siguiente sistema de ecuaciones diferenciales parciales*

$$\begin{aligned}
&\bar{n}_t(t,\boldsymbol{p})-\bar{n}(t,\boldsymbol{p})(\lambda_1+\lambda_2)+\sum_{i,j\in\mathcal{E}}\bar{n}_{p_i}(t,\boldsymbol{p})q^{ij}p^j+2\bar{m}(t)(\Gamma_1+\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p}+\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})\\
&\quad+\frac{1}{2}\sum_{i,j\in\mathcal{E}}\bar{n}_{p_ip_j}(t,\boldsymbol{p})(H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p})+H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p}))-\Theta_2(\boldsymbol{p})=0, \qquad (4.21)\\
&\bar{u}_t(t,\boldsymbol{p})+\Gamma_2\bar{m}(t)+\bar{n}(t,\boldsymbol{p})(\Gamma_1+\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p}+\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})+\Theta_3(\boldsymbol{p})
\end{aligned}$$

$$+\sum_{i,j\in\mathcal{E}}\bar{u}_{p_i}(t,\boldsymbol{p})q^{ij}p^j+\frac{1}{2}\sum_{i,j\in\mathcal{E}}\bar{u}_{p_ip_j}(t,\boldsymbol{p})(H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p})+H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p}))$$

$$+\sum_{i\in\mathcal{E}}\bar{n}_{p_i}(t,\boldsymbol{p})\left[\frac{b_1^2}{\sqrt{\sigma^2+b_2^2}}H^{i,(1)}(\boldsymbol{p})-\frac{\sqrt{\sigma^2(b_1^2+b_2^2)+b_2^2b_1^2}}{\sqrt{\sigma^2+b_1^2}}H^{i,(2)}(\boldsymbol{p})\right]=0, \tag{4.22}$$

con condiciones terminales $\bar{n}(T,\boldsymbol{p})=0$ *y* $\bar{u}(T,\boldsymbol{p})=0$ *y donde los parámetros* Γ_1 *y* Γ_2 *están dados como en el Teorema 3.1.1 y*

$$\begin{aligned}
\Theta_1&=\frac{(\sigma^2+b_2^2)\lambda_1^2+(\sigma^2+b_1^2)\lambda_2^2+2\sigma^2\lambda_1\lambda_2}{2(\sigma^2(b_1^2+b_2^2)+b_1^2b_2^2)},\\
\Theta_2(\boldsymbol{p})&=\frac{\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p}(\lambda_1(b_2^2+\sigma^2)+\lambda_2\sigma^2)+\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p}(\lambda_2(b_1^2+\sigma^2)+\lambda_1\sigma^2)}{b_1^2b_2^2+\sigma^2(b_1^2+b_2^2)},\\
\Theta_3(\boldsymbol{p})&=r+\frac{\mu_m^2}{2\sigma_m^2}+\frac{(\lambda_1b_2\boldsymbol{\alpha}_1^T\boldsymbol{p})^2+(\lambda_2b_1\boldsymbol{\alpha}_2^T\boldsymbol{p})^2+\sigma^2(\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p}+\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})^2}{2(b_1^2b_2^2+\sigma^2(b_1^2+b_2^2))}.
\end{aligned}$$

En lo anterior $\bar{m}_t,\bar{n}_t$, $\bar{u}_t$, $\bar{n}_{p_i}$ *y* $\bar{u}_{p_i}$ *denotan las derivadas parciales de las funciones* $\bar{m}$, $\bar{n}$ *y* $\bar{u}$, *respectivamente.*

Demostración. A partir de la ecuación de Hamilton-Jacobi-Bellman se obtendrá la estrategia óptima $h_*=(h_*^{(m)},h_*^{(1)},h_*^{(2)})$ para luego proponer un candidato a la función de valor y finalmente verificar que éste coincida con la solución de la ecuación HJB.

1. Estrategia Óptima

Supongamos que la función de valor V es de clase $\mathcal{C}^{1,2,2,2}([0,T],\mathbb{R}^+,\mathbb{R},\Delta_k)$, lo cual implica que satisface la ecuación de Hamilton-Jacobi-Bellman, dada por (3.4), sujeto a la condición terminal $V(T,\cdot)=\log(w)$.

Por otro lado, por (4.18) podemos escribir a la función de valor como

$$\begin{aligned}
V(t,w,x,\boldsymbol{p})&=\sup_{h\in\mathcal{A}^{\mathbb{F}}}\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_T^h]\\
&=\sup_{h\in\mathcal{A}^{\mathbb{F}}}\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log(W_t^h\exp\{G(t,T,X_s,\boldsymbol{\pi}_s;t\le s\le T)\})]\\
&=\sup_{h\in\mathcal{A}^{\mathbb{F}}}\left\{\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_t^h]+\mathbb{E}_{t,w,x,\boldsymbol{p}}[G(t,T,X_s,\boldsymbol{\pi}_s;t\le s\le T)]\right\}\\
&=\log(w)+\sup_{h\in\mathcal{A}^{\mathbb{F}}}\mathbb{E}_{t,w,x,\boldsymbol{p}}[G(t,T,X_s,\boldsymbol{\pi}_s;t\le s\le T)]\\
&=:\log(w)+v(t,x,\boldsymbol{p}),
\end{aligned}\tag{4.23}$$

para alguna función $v(t,x,\boldsymbol{p})$, con $v(T,\cdot,\cdot)=0$.

Entonces, sustituyendo en (4.19), la ecuación de Hamilton-Jacobi-Bellman es de la siguiente forma

$$\begin{aligned}
\sup_{h\in\mathcal{A}^{\mathbb{F}}} \Bigg\{ & v_t(t,x,\boldsymbol{p}) + v_x(t,x,\boldsymbol{p})[\Gamma_1 + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}] + \sum_{i,j\in\mathcal{E}} v_{p_i}(t,x,\boldsymbol{p})q^{ij}p^j \\
& + r + (h^{(m)} + h^{(1)}\beta_1 + h^{(2)}\beta_2)\mu_m + h^{(1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} + h^{(2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p} + \frac{1}{2}\Gamma_2 v_{xx}(t,x,\boldsymbol{p}) \\
& + \frac{1}{2}\sum_{i,j\in\mathcal{E}} v_{p_ip_j}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big] \\
& - \frac{1}{2}\big[(h^{(m)} + \beta_1 h^{(1)} + \beta_2 h^{(2)})^2\sigma_m^2 + (\sigma_1 h^{(1)} + \rho\sigma_2 h^{(2)})^2 + \sigma_2^2(1-\rho^2)(h^{(2)})^2\big] \\
& + \sum_{i\in\mathcal{E}} v_{xp_i}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1 - \rho\sigma_2) - H^{i,(2)}(\boldsymbol{p})\sigma_2\sqrt{1-\rho^2}\big]\Bigg\} = 0. \qquad (4.24)
\end{aligned}$$

Para hallar el supremo de la ecuación anterior, mediante las condiciones de primer orden se obtendrán los puntos críticos. Luego, con las condiciones de segundo orden se comprobará que son máximos, obteniendo así la estrategia óptima $h_* = (h_*^{(m)}, h_*^{(1)}, h_*^{(2)})$.

De las condiciones de primer orden, derivando (4.24) respecto de $h^{(m)}$, $h^{(1)}$ y $h^{(2)}$ obtenemos, respectivamente

$$\frac{\mu_m}{\sigma_m^2} - (h_*^{(m)} + \beta_1 h_*^{(1)} + \beta_2 h_*^{(2)}) = 0, \qquad (4.25)$$

$$\beta_1\mu_m + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \beta_1\sigma_m^2(h_*^{(m)} + \beta_1 h_*^{(1)} + \beta_2 h_*^{(2)}) - (\sigma_1^2 h_*^{(1)} + \rho\sigma_1\sigma_2 h_*^{(2)}) = 0, \qquad (4.26)$$

$$\mu_m\beta_2 + \boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \beta_2\sigma_m^2(h_*^{(m)} + \beta_1 h_*^{(1)} + \beta_2 h_*^{(2)}) - (\rho^2\sigma_2^2 h_*^{(2)} + \rho\sigma_1\sigma_2 h_*^{(1)}) - \sigma_2^2(1-\rho^2)h_*^{(2)} = 0. \qquad (4.27)$$

De donde,

$$h_*^{(m)} = \frac{\mu_m}{\sigma_m^2} - \beta_1 h_*^{(1)} - \beta_2 h_*^{(2)}, \qquad (4.28)$$

sustituyendo (4.28) en (4.27),

$$\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - (\sigma_2^2\rho^2 h_*^{(2)} + \rho\sigma_1\sigma_2 h_*^{(1)}) - \sigma_2^2(1-\rho^2)h_*^{(2)} = 0,$$

esto es,

$$\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \sigma_2^2 h_*^{(2)} - \rho\sigma_1\sigma_2 h_*^{(1)} = 0,$$

por lo que

$$h_*^{(2)} = \frac{\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \sigma^2 h_*^{(1)}}{\sigma_2^2}, \qquad (4.29)$$

ya que $\rho = \frac{\sigma^2}{\sigma_1\sigma_2}$.

Luego, sustituyendo (4.28) y (4.29) en (4.26) se tiene que

$$\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \sigma_1^2 h_*^{(1)} - \sigma^2\left(\frac{\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \sigma^2 h_*^{(1)}}{\sigma_2^2}\right) = 0,$$

entonces,

$$\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \frac{\boldsymbol{\mu}_2(x)^T\boldsymbol{p}\sigma^2}{\sigma^2 + b_2^2} - h_*^{(1)}\left(\sigma^2 + b_1^2 - \sigma^2 \cdot \frac{\sigma^2}{\sigma^2 + b_2^2}\right) = 0,$$

pues $\sigma_1 = \sqrt{\sigma^2 + b_1^2}$ y $\sigma_2 = \sqrt{\sigma^2 + b_2^2}$.

Por lo tanto,

$$h_*^{(1)} = \frac{\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}\varrho_2}{b_1^2 + b_2^2\varrho_2}, \tag{4.30}$$

con $\varrho_2 = \frac{\sigma^2}{\sigma^2 + b_2^2}$ y donde se utilizó que

$$\sigma^2 + b_1^2 - \sigma^2 \cdot \frac{\sigma^2}{\sigma^2 + b_2^2} = \frac{\sigma^2(b_1^2 + b_2^2) + b_1^2 b_2^2}{\sigma^2 + b_2^2} = \frac{b_1^2(\sigma^2 + b_2^2) + b_2^2\sigma^2}{\sigma^2 + b_2^2} = b_1^2 + b_2^2\varrho_2.$$

Finalmente, de (4.29) y (4.30), obtenemos

$$\begin{aligned} h_*^{(2)} &= \frac{\boldsymbol{\mu}_2(x)^T\boldsymbol{p}}{\sigma^2 + b_2^2} - \varrho_2\left[\frac{\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}\varrho_2}{b_1^2 + b_2^2\varrho_2}\right] \\ &= \boldsymbol{\mu}_2(x)^T\boldsymbol{p}\left[\frac{1}{\sigma^2 + b_2^2} + \frac{\varrho_2^2}{b_1^2 + b_2^2\varrho_2}\right] - \frac{\boldsymbol{\mu}_1(x)^T\boldsymbol{p}\varrho_2}{b_1^2 + b_2^2\varrho_2}. \end{aligned} \tag{4.31}$$

Observemos que

$$\frac{\varrho_2}{b_1^2 + b_2^2\varrho_2} = \frac{\sigma^2}{b_1^2(\sigma^2 + b_2^2) + \sigma^2 b_2^2} = \frac{\sigma^2}{b_2^2(b_1^2 + \sigma^2) + b_1^2\sigma^2} = \frac{\varrho_1}{b_2^2 + b_1^2\varrho_1}, \tag{4.32}$$

con $\varrho_1 = \frac{\sigma^2}{\sigma^2 + b_1^2}$.

Por otro lado,

$$\begin{aligned} \frac{1}{\sigma^2 + b_2^2} + \frac{\varrho^2}{b_1^2 + b_2^2\varrho_2} &= \frac{b_1^2 + b_2^2\varrho_2 + \varrho^2(\sigma^2 + b_2^2)}{(\sigma^2 + b_2^2)(b_1^2 + b_2^2\varrho_2)} \\ &= \frac{b_1^2 + b_2^2\varrho_2 + \varrho^2(\sigma^2 + b_2^2)}{b_1^2(\sigma^2 + b_2^2) + b_2^2\sigma^2} \\ &= \frac{b_1^2 + b_2^2\varrho_2 + \varrho^2(\sigma^2 + b_2^2)}{b_2^2(\sigma^2 + b_1^2) + b_1^2\sigma^2} \end{aligned}$$

$$
\begin{aligned}
&= \frac{b_1^2 + b_2^2\varrho_2 + \varrho^2(\sigma^2 + b_2^2)}{\sigma^2 + b_1^2} \cdot \frac{1}{b_2^2 + b_1^2\varrho_1} \\
&= \frac{b_1^2(\sigma^2 + b_2^2) + b_2^2\sigma^2 + \sigma^4}{(\sigma^2 + b_1^2)(\sigma^2 + b_2^2)} \cdot \frac{1}{b_2^2 + b_1^2\varrho_1} \\
&= \frac{1}{b_2^2 + b_1^2\varrho_1}. \qquad (4.33)
\end{aligned}
$$

Así, por (4.32) y (4.33), de (4.31) se sigue que

$$
h_*^{(2)} = \frac{\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \boldsymbol{\mu}_1(x)^T\boldsymbol{p}\varrho_1}{b_2^2 + b_1^2\varrho_2}.
$$

Veamos ahora que $h_* = (h_*^{(m)}, h_*^{(1)}, h_*^{(2)})$ en efecto es óptimo.

Obteniendo las segundas derivadas se llega a

$$
A = \begin{pmatrix} -\sigma_m^2 & -\beta_1\sigma_m^2 & -\beta_2\sigma_m^2 \\ -\beta_1\sigma_m^2 & -\beta_1^2\sigma_m^2 - \sigma_1^2 & -\beta_1\beta_2\sigma_m^2 - \sigma^2 \\ -\beta_2\sigma_m^2 & -\beta_1\beta_2\sigma_m^2 - \sigma^2 & -\beta_2^2\sigma_m^2 - \sigma_2^2 \end{pmatrix}.
$$

Notemos que tal matriz es igual a la del Teorema 3.1.1, por tanto es definida no positiva y así concluimos que $h_* = (h_*^{(m)}, h_*^{(1)}, h_*^{(2)})$ es la estrategia óptima.

2. Existencia

Caracterizaremos ahora la forma de la función de valor. Para ello sustituiremos $h_* = (h_*^{(m)}, h_*^{(1)}, h_*^{(2)})$ en (4.24).

Notemos primero que de las ecuaciones (4.26), (4.27) y (4.28), se tiene que

$$
\begin{aligned}
\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \sigma_1^2 h_*^{(1)} - \rho\sigma_1\sigma_2 h_*^{(2)} = 0, \\
\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \sigma_2^2 h_*^{(2)} - \rho\sigma_1\sigma_2 h_*^{(1)} = 0,
\end{aligned}
$$

multiplicando las ecuaciones anteriores por $h_*^{(1)}$ y $h_*^{(2)}$, respectivamente

$$
\begin{aligned}
\boldsymbol{\mu}_1(x)^T\boldsymbol{p}h_*^{(1)} - \sigma_1^2(h_*^{(1)})^2 - \rho\sigma_1\sigma_2 h_*^{(2)}h_*^{(1)} = 0, \\
\boldsymbol{\mu}_2(x)^T\boldsymbol{p}h_*^{(2)} - \sigma_2^2(h_*^{(2)})^2 - \rho\sigma_1\sigma_2 h_*^{(1)}h_*^{(2)} = 0,
\end{aligned}
$$

esto es,

$$
\boldsymbol{\mu}_1(x)^T\boldsymbol{p}h_*^{(1)} - \sigma_1^2(h_*^{(1)})^2 - 2\rho\sigma_1\sigma_2 h_*^{(2)}h_*^{(1)} + \boldsymbol{\mu}_2(x)^T\boldsymbol{p}h_*^{(2)} - \sigma_2^2(h_*^{(2)})^2 = 0. \qquad (4.34)
$$

Entonces, reemplazando $h_* = (h_*^{(m)}, h_*^{(1)}, h_*^{(2)})$ en (4.24),

$$\begin{aligned}
v_t(t,x,\boldsymbol{p}) &+ v_x(t,x,\boldsymbol{p})[\Gamma_1 + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}] + \sum_{i,j\in\mathcal{E}} v_{p_i}(t,x,\boldsymbol{p})q^{ij}p^j \\
&+ r + \frac{\mu_m^2}{2\sigma_m^2} + h_*^{(1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} + h_*^{(2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p} + \frac{1}{2}\Gamma_2 v_{xx}(t,x,\boldsymbol{p}) \\
&+ \frac{1}{2}\sum_{i,j\in\mathcal{E}} v_{p_ip_j}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big] \\
&- \frac{1}{2}\big[(\sigma_1 h_*^{(1)})^2 + 2\rho\sigma_1\sigma_2 h_*^{(1)}h_*^{(2)} + (\sigma_2 h_*^{(2)})^2\big] \\
&+ \sum_{i\in\mathcal{E}} v_{x,p_i}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1 - \rho\sigma_2) - H^{i,(2)}(\boldsymbol{p})\sigma_2\sqrt{1-\rho^2}\big] = 0,
\end{aligned}$$

y por (4.34),

$$\begin{aligned}
v_t(t,x,\boldsymbol{p}) &+ v_x(t,x,\boldsymbol{p})[\Gamma_1 + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}] + \sum_{i,j\in\mathcal{E}} v_{p_i}(t,x,\boldsymbol{p})q^{ij}p^j \\
&+ r + \frac{\mu_m^2}{2\sigma_m^2} + \frac{1}{2}\big[h_*^{(1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} + h_*^{(2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p}\big] + \frac{1}{2}\Gamma_2 v_{xx}(t,x,\boldsymbol{p}) \\
&+ \frac{1}{2}\sum_{i,j\in\mathcal{E}} v_{p_ip_j}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big] \\
&+ \sum_{i\in\mathcal{E}} v_{x,p_i}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1 - \rho\sigma_2) - H^{i,(2)}(\boldsymbol{p})\sigma_2\sqrt{1-\rho^2}\big] = 0.
\end{aligned}$$

Por otro lado, notemos que

$$\boldsymbol{\mu}_1(X_t)^T\boldsymbol{\pi}_t = -\lambda_1\mathbb{E}(X_t - \alpha_1(Y_t)|\mathcal{F}_t) = -\lambda_1(X_t - \boldsymbol{\alpha}_1^T\boldsymbol{\pi}_t)$$

y

$$\boldsymbol{\mu}_2(X_t)^T\boldsymbol{\pi}_t = \lambda_2\mathbb{E}(X_t - \alpha_2(Y_t)|\mathcal{F}_t) = \lambda_2(X_t - \boldsymbol{\alpha}_2^T\boldsymbol{\pi}_t).$$

Por lo que,

$$\begin{aligned}
\boldsymbol{\mu}_1(x)^T\boldsymbol{p}h_*^{(1)} &= \frac{\lambda_1^2(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})(x-\boldsymbol{\alpha}_1^T\boldsymbol{p}) + \lambda_1\lambda_2(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})\varrho_2}{b_1^2 + b_2^2\varrho_2} \\
&= \frac{\lambda_1^2(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})(\sigma^2+b_2^2) + \lambda_1\lambda_2(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})\sigma^2}{\sigma^2(b_1^2+b_2^2) + b_1^2b_2^2} \\
&= \frac{\lambda_1^2(\sigma^2+b_2^2)(x^2 - 2x\boldsymbol{\alpha}_1^T\boldsymbol{p} + (\boldsymbol{\alpha}_1^T\boldsymbol{p})^2) + \lambda_1\lambda_2\sigma^2(x^2 - x[\boldsymbol{\alpha}_1^T\boldsymbol{p} + \boldsymbol{\alpha}_2^T\boldsymbol{p}] + \boldsymbol{\alpha}_1^T\boldsymbol{p}\boldsymbol{\alpha}_2^T\boldsymbol{p})}{\sigma^2(b_1^2+b_2^2) + b_1^2b_2^2} \\
&= \frac{x^2(\lambda_1^2(\sigma^2+b_2^2) + \lambda_1\lambda_2\sigma^2) - x(\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p}(\lambda_2\sigma^2 + 2\lambda_1(\sigma^2+b_2^2)) + \lambda_1\lambda_2\sigma^2\boldsymbol{\alpha}_2^T\boldsymbol{p})}{\sigma^2(b_1^2+b_2^2) + b_1^2b_2^2} \\
&\quad + \frac{\sigma^2(\lambda_1^2(\boldsymbol{\alpha}_1^T\boldsymbol{p})^2 + \lambda_1\lambda_2\boldsymbol{\alpha}_1^T\boldsymbol{p}\boldsymbol{\alpha}_2^T\boldsymbol{p}) + (\lambda_1 b_2\boldsymbol{\alpha}_1^T\boldsymbol{p})^2}{\sigma^2(b_1^2+b_2^2) + b_1^2b_2^2},
\end{aligned}$$

y

$$
\begin{aligned}
\boldsymbol{\mu}_2(x)^T \boldsymbol{p} h_*^{(2)} &= \frac{\lambda_2^2(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})+\lambda_1\lambda_2(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})\varrho_1}{b_2^2+b_1^2\varrho_1}\\
&= \frac{\lambda_2^2(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})(\sigma^2+b_1^2)+\lambda_1\lambda_2\sigma^2(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})}{\sigma^2(b_1^2+b_2^2)+b_1^2b_2^2}\\
&= \frac{\lambda_2^2(\sigma^2+b_1^2)(x^2-2x\boldsymbol{\alpha}_2^T\boldsymbol{p}+(\boldsymbol{\alpha}_2^T\boldsymbol{p})^2)+\lambda_1\lambda_2\sigma^2(x^2-x[\boldsymbol{\alpha}_1^T\boldsymbol{p}+\boldsymbol{\alpha}_2^T\boldsymbol{p}]+\boldsymbol{\alpha}_1^T\boldsymbol{p}\boldsymbol{\alpha}_2^T\boldsymbol{p})}{\sigma^2(b_1^2+b_2^2)+b_1^2b_2^2}\\
&= \frac{x^2(\lambda_2^2(\sigma^2+b_1^2)+\lambda_1\lambda_2\sigma^2)-x(\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p}(\lambda_1\sigma^2+2\lambda_2(\sigma^2+b_1^2))+\lambda_1\lambda_2\sigma^2\boldsymbol{\alpha}_1^T\boldsymbol{p})}{\sigma^2(b_1^2+b_2^2)+b_1^2b_2^2}\\
&\quad+\frac{\sigma^2((\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})^2+\lambda_1\lambda_2\boldsymbol{\alpha}_1^T\boldsymbol{p}\boldsymbol{\alpha}_2^T\boldsymbol{p})+(\lambda_2 b_1\boldsymbol{\alpha}_2^T\boldsymbol{p})^2}{\sigma^2(b_1^2+b_2^2)+b_1^2b_2^2}.
\end{aligned}
$$

Entonces,

$$
\begin{aligned}
&v_t(t,x,\boldsymbol{p})+v_x(t,x,\boldsymbol{p})[\Gamma_1-\lambda_1(x-\alpha_1^T\boldsymbol{p})-\lambda_2(x-\alpha_2^T\boldsymbol{p})]+\sum_{i,j\in\mathcal{E}} v_{p_i}(t,x,\boldsymbol{p})q^{ij}p^j\\
&\quad+\frac{1}{2}\sum_{i,j\in\mathcal{E}} v_{p_ip_j}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p})+H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big]\\
&\quad+\sum_{i\in\mathcal{E}} v_{xp_i}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1-\rho\sigma_2)-H^{i,(2)}(\boldsymbol{p})\sigma_2\sqrt{1-\rho^2}\big]\\
&\quad+\frac{1}{2}\Gamma_2 v_{xx}(t,x,\boldsymbol{p})+\Theta_1x^2-\Theta_2(\boldsymbol{p})x+\Theta_3(\boldsymbol{p})=0,
\end{aligned}
\tag{4.35}
$$

con $\Theta_1, \Theta_2(\boldsymbol{p})$ y $\Theta_3(\boldsymbol{p})$ como en el enunciado del teorema.

Ahora bien, proponemos como función de valor $v(t,x,\boldsymbol{p})=\bar{m}(t)x^2+\bar{n}(t,\boldsymbol{p})x+\bar{u}(t,\boldsymbol{p})$. Si sustituimos en (4.35), se satisface

$$
\begin{aligned}
&\bar{m}_t(t)x^2+\bar{n}_t(t,\boldsymbol{p})x+\bar{u}_t(t,\boldsymbol{p})+[\Gamma_1-\lambda_1(x-\alpha_1^T\boldsymbol{p})-\lambda_2(x-\alpha_2^T\boldsymbol{p})][2\bar{m}(t)x+\bar{n}(t,\boldsymbol{p})]\\
&\quad+\frac{1}{2}\sum_{i,j\in\mathcal{E}}\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p})+H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big][\bar{n}_{p_ip_j}(t,\boldsymbol{p})x+\bar{u}_{p_ip_j}(t,\boldsymbol{p})]\\
&\quad+\sum_{i\in\mathcal{E}}\bar{n}_{p_i}(t,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1-\rho\sigma_2)-H^{i,(2)}(\boldsymbol{p})\sigma_2\sqrt{1-\rho^2}\big]+\Gamma_2\bar{m}(t)\\
&\quad+\Theta_1x^2-\Theta_2(\boldsymbol{p})x+\Theta_3(\boldsymbol{p})+\sum_{i,j\in\mathcal{E}} q^{ij}p^j\big[\bar{n}_{p_i}(t,\boldsymbol{p})x+\bar{u}_{p_i}(t,\boldsymbol{p})\big]=0,
\end{aligned}
$$

es decir,

$$
\begin{aligned}
&x^2\big[\bar{m}_t(t)-2\bar{m}(t)(\lambda_1+\lambda_2)+\Theta_1\big]\\
&\quad+x\Bigg[\bar{n}_t(t,\boldsymbol{p})-\bar{n}(t,\boldsymbol{p})(\lambda_1+\lambda_2)+\sum_{i,j\in\mathcal{E}}\bar{n}_{p_i}(t,\boldsymbol{p})q^{ij}p^j+2\bar{m}(t)(\Gamma_1+\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p}+\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2}\sum_{i,j\in\mathcal{E}}\bar{n}_{p_ip_j}(t,\boldsymbol{p})(H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p})+H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p}))-\Theta_2(\boldsymbol{p})\Big]\\
&+\bar{u}_t(t,\boldsymbol{p})+\Gamma_2\bar{m}(t)+\bar{n}(t,\boldsymbol{p})(\Gamma_1+\lambda_1\boldsymbol{\alpha}_1^t\boldsymbol{p}+\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})+\Theta_3(\boldsymbol{p})\\
&+\sum_{i,j\in\mathcal{E}}\bar{u}_{p_i}(t,\boldsymbol{p})q^{ij}p^j+\frac{1}{2}\sum_{i,j\in\mathcal{E}}\bar{u}_{p_ip_j}(t,\boldsymbol{p})(H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p})+H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p}))\\
&+\sum_{i\in\mathcal{E}}\bar{n}_{p_i}(t,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1-\rho\sigma_2)-H^{i,(2)}(\boldsymbol{p})\sigma_2\sqrt{1-\rho^2}\big]=0.
\end{aligned}
$$

Notemos que,

$$
\begin{aligned}
\sigma_1-\rho\sigma_2&=\frac{b_1^2}{\sqrt{\sigma^2+b_1^2}},\\
\sigma_2\sqrt{1-\rho^2}&=\sqrt{\sigma_2^2-\sigma_2^2\rho^2}=\frac{\sqrt{\sigma^2(b_1^2+b_2^2)+b_1^2b_2^2}}{\sqrt{\sigma^2+b_1^2}}.
\end{aligned}
$$

Por lo tanto, si las funciones $\bar{m}(t)$, $\bar{n}(t,\boldsymbol{p})$ y $\bar{u}(t,\boldsymbol{p})$ satisfacen (4.20), (4.21) y (4.22), con condiciones terminales $\bar{m}(T)=0$, $\bar{n}(T,\boldsymbol{p})=0$ y $\bar{u}(T,\boldsymbol{p})=0$ entonces, $v(t,x,\boldsymbol{p})=\bar{m}(t)x^2+\bar{n}(t,\boldsymbol{p})x+\bar{u}(t,\boldsymbol{p})$ resuelve la ecuación de HJB y así, la función de valor es de la forma

$$
V(t,w,x,\boldsymbol{p})=\log(w)+v(t,x,\boldsymbol{p})=\log(w)+\bar{m}(t)x^2+\bar{n}(t,\boldsymbol{p})x+\bar{u}(t,\boldsymbol{p}).
$$

3. Verificación

Sea $\nu(t,w,x,\boldsymbol{p})$ de clase $C^{1,2,2,2}([0,T],\mathbb{R}^+,\mathbb{R},\Delta_k)$ y con crecimiento cuadrático una solución a la ecuación HJB, dada en (4.24), y $h\in\mathcal{A}^{\mathbb{F}}$ una estrategia admisible.

Para cada $n\geq 1$ definimos los siguientes tiempos de paro

$$
\tau_n := T\wedge \text{ínf}\{s>t : \|(W_s,X_s,\boldsymbol{\pi}_s)-(w,x,\boldsymbol{p})\|\geq n\}.
$$

De la definición se observa que τ_n converge a T $\mathbb{P}$-c.s.

Aplicando la fórmula de Itô a $\nu(\tau_n,W_{\tau_n},X_{\tau_n},\boldsymbol{\pi}_{\tau_n})$ tenemos que

$$
\begin{aligned}
d\nu(\tau_n,W_{\tau_n},X_{\tau_n},\boldsymbol{\pi}_{\tau_n})=&\nu_t(t,W_t^h,X_t,\boldsymbol{\pi}_t)dt+\nu_w(t,W_t^h,X_t,\boldsymbol{\pi}_t)dW_t^h+\nu_x(t,W_t^h,X_t,\boldsymbol{\pi}_t)dX_t\\
&+\sum_{i\in\mathcal{E}}\nu_{p_i}(t,W_t^h,X_t,\boldsymbol{\pi}_t)d\pi_t^i+\frac{1}{2}\nu_{wx}(t,W_t^h,X_t,\boldsymbol{\pi}_t)d[W,X]_t\\
&+\frac{1}{2}\nu_{ww}(t,W_t^h,X_t,\boldsymbol{\pi}_t)d[W,W]_t+\frac{1}{2}\nu_{xx}(t,W_t^h,X_t,\boldsymbol{\pi}_t)d[X,X]_t\\
&+\frac{1}{2}\sum_{i\in\mathcal{E}}\nu_{wp_i}(t,W_t^h,X_t,\boldsymbol{\pi}_t)d[W,\pi^i]_t+\frac{1}{2}\sum_{i\in\mathcal{E}}\nu_{xp_i}(t,W_t^h,X_t,\boldsymbol{\pi}_t)d[X,\pi^i]_t
\end{aligned}
$$

$$+\frac{1}{2}\sum_{j,i\in\mathcal{E}}\nu_{p_ip_j}(t,W_t^h,X_t,\boldsymbol{\pi}_t)d[\pi^i,\pi^j]. \tag{4.36}$$

Usando la independencia entre los procesos I y $B^{(m)}$, de la Proposición 4.1.1 y las ecuaciones (4.16) y (4.17), obtenemos que

$$\begin{aligned}
d[W,W]_t &= W_t^2\big[\sigma_m^2(h^{(m)}+h^{(1)}\beta_1+h^{(2)}\beta_2)^2dt+(\sigma_1h^{(1)}+\rho\sigma_2h^{(2)})^2dt+\sigma_2^2(1-\rho^2)h^{(2)}dt\big]\\
d[W,X]_t &= W_t\big[\sigma_m^2(\beta_1-\beta_2)(h^{(m)}+h^{(1)}\beta_1+h^{(2)}\beta_2)dt+(\sigma_1^2h^{(1)}-\rho\sigma_1\sigma_2h^{(1)}-\sigma_2^2h^{(2)}\\
&\qquad+\sigma_1\sigma_2\rho h^{(2)})dt\big]\\
d[X,X]_t &= (\sigma_m^2(\beta_1-\beta_2)^2+b_1^2+b_2^2)dt\\
d[W,\pi^i]_t &= W_t\big[H^{i,(1)}(\boldsymbol{\pi}_t)(\sigma_1h^{(1)}+\sigma_2\rho h^{(2)})dt+H^{i,(2)}(\boldsymbol{\pi}_t)\big(\sigma_2\sqrt{1-\rho^2}h^{(2)}\big)dt\big]\\
d[X,\pi^i]_t &= H^{i,(1)}(\boldsymbol{\pi}_t)(\sigma_1-\sigma_2\rho)dt-H^{i,(2)}(\boldsymbol{\pi}_t)\sigma_2\sqrt{1-\rho^2}dt\\
d[\pi^i,\pi^j]_t &= H^{i,(1)}(\boldsymbol{\pi}_t)H^{j,(1)}(\boldsymbol{\pi}_t)dt+H^{i,(2)}(\boldsymbol{\pi}_t)H^{j,(2)}(\boldsymbol{\pi}_t)dt.
\end{aligned} \tag{4.37}$$

Por lo tanto, de forma integral se tiene que

$$\begin{aligned}
\nu(\tau_n,W_{\tau_n}^h,X_{\tau_n},\boldsymbol{\pi}_{\tau_n}) &= \nu(t,w,x,\boldsymbol{p})+\int_t^{\tau_n}\mathcal{L}_{\mathbb{F}}^h\nu(s,W_s^h,X_s,\boldsymbol{\pi}_s)ds\\
&+\int_t^{\tau_n}\big[W_s^h\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)(h_s^{(m)}+\beta_1h_s^{(1)}+\beta_2h_s^{(2)})+\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\beta_1-\beta_2)\big]\sigma_mdB_s^{(m)}\\
&+\int_t^{\tau_n}\big[W_s^h\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\sigma_1h_s^{(1)}+\rho\sigma_2h_s^{(2)})+\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\sigma_1-\rho\sigma_2)\big]dI_s^{(1)}\\
&+\int_t^{\tau_n}\sum_{i\in\mathcal{E}}\nu_{p_i}(s,W_s^h,X_s,\boldsymbol{\pi}_s)H^{i,(1)}(\boldsymbol{\pi}_s)dI_s^{(1)}\\
&+\int_t^{\tau_n}\sigma_2\sqrt{1-\rho^2}\big[W_s^hh_s^{(2)}\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)-\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)\big]dI_s^{(2)}\\
&+\int_t^{\tau_n}\sum_{i\in\mathcal{E}}\nu_{p_i}(s,W_s^h,X_s,\boldsymbol{\pi}_s)H^{i,(2)}(\boldsymbol{\pi}_s)dI_s^{(2)},
\end{aligned} \tag{4.38}$$

donde $\mathcal{L}_{\mathbb{F}}^h$ denota el generador del proceso $(W,X,\boldsymbol{\pi})$ definido en (4.19).

Además, como ν satisface la ecuación HJB, entonces de (4.38) tenemos

$$\begin{aligned}
\nu(\tau_n,W_{\tau_n}^h,X_{\tau_n},\boldsymbol{\pi}_{\tau_n}) &\le \nu(t,w,x,\boldsymbol{p})+\int_t^{\tau_n}\sum_{i\in\mathcal{E}}\nu_{p_i}(s,W_s^h,X_s,\boldsymbol{\pi}_s)H^{i,(1)}(\boldsymbol{\pi}_s)dI_s^{(1)}\\
&+\int_t^{\tau_n}\big[W_s^h\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)(h_s^{(m)}+\beta_1h_s^{(1)}+\beta_2h_s^{(2)})+\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\beta_1-\beta_2)\big]\sigma_mdB_s^{(m)}\\
&+\int_t^{\tau_n}\big[W_s^h\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\sigma_1h_s^{(1)}+\rho\sigma_2h_s^{(2)})+\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\sigma_1-\rho\sigma_2)\big]dI_s^{(1)}\\
&+\int_t^{\tau_n}\sigma_2\sqrt{1-\rho^2}\big[W_s^hh_s^{(2)}\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)-\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)\big]dI_s^{(2)}
\end{aligned}$$

$$+\int_t^{\tau_n}\sum_{i\in\mathcal{E}}\nu_{p_i}(s,W_s^h,X_s,\boldsymbol{\pi}_s)H^{i,(2)}(\boldsymbol{\pi}_s)dI_s^{(2)}.$$

Luego, por definición de τ_n, por ser ν función de clase $C^{1,2,2,2}([0,T],\mathbb{R}^+,\mathbb{R},\Delta_k)$ y al cumplir h la condición de integrabilidad (3.1), podemos concluir que los integrandos respecto a $B^{(m)}$, $I^{(1)}$ y $I^{(2)}$ son acotados y así tales integrales tienen esperanza cero.

Por tanto, tomando esperanza en la ecuación anterior se sigue que $\mathbb{E}_{t,w,x,\boldsymbol{p}}[\nu(\tau_n,W_{\tau_n}^h,X_{\tau_n},\boldsymbol{\pi}_{\tau_n})]\leq\nu(t,w,x,\boldsymbol{p})$.

Por otro lado, al ser ν continua se tiene que

$$\mathbb{E}_{t,w,x,\boldsymbol{p}}[\nu(T,W_T^h,X_T,\boldsymbol{\pi}_T)]=\mathbb{E}_{t,w,x,\boldsymbol{p}}[\lim_{n\to\infty}\nu(\tau_n,W_{\tau_n}^h,X_{\tau_n},\boldsymbol{\pi}_{\tau_n})],\quad \mathbb{P}\text{-c.s.}$$

Luego, como ν tiene crecimiento cuadrático, se satisface que

$$|\nu(\tau_n,W_{\tau_n}^h,X_{\tau_n},\boldsymbol{\pi}_{\tau_n})|\leq C(1+\sup_{s\leq T}\{|W_s^h|^2+|X_s|^2+|\boldsymbol{\pi}_s|^2\}),$$

y por el Teorema 1.1.6, al ser $\alpha(\cdot)$ localmente Lipschitz, continua y con crecimiento lineal, se tiene que $\sup_{s\leq T}\{|W_s^h|^2+|X_s|^2+|\boldsymbol{\pi}_s|^2\}$ es integrable. Entonces, por convergencia dominada obtenemos

$$\begin{aligned}\mathbb{E}_{t,w,x,\boldsymbol{p}}[\nu(T,W_T^h,X_T,\boldsymbol{\pi}_T)]&=\mathbb{E}_{t,w,x,\boldsymbol{p}}[\lim_{n\to\infty}\nu(\tau_n,W_{\tau_n}^h,X_{\tau_n},\boldsymbol{\pi}_{\tau_n})]\\&=\lim_{n\to\infty}\mathbb{E}_{t,w,x,\boldsymbol{p}}[\nu(\tau_n,W_{\tau_n}^h,X_{\tau_n},\boldsymbol{\pi}_{\tau_n})]\\&\leq\nu(t,w,x,\boldsymbol{p}).\end{aligned}$$

Utilizando la condición terminal $\nu(T,\cdot)=\log(\cdot)$, obtenemos $\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_t^h]\leq\nu(t,w,x,\boldsymbol{p})$ y tomando supremo sobre las estrategias admisibles llegamos a $V(t,w,x,\boldsymbol{p})\leq\nu(t,w,x,\boldsymbol{p})$.

Finalmente, bajo el mismo razonamiento anterior, ahora si $h_*=(h_*^{(m)},h_*^{(1)},h_*^{(2)})$ es la estrategia óptima, al tomar esperanza en la ecuación (4.38) se sigue que

$$\nu(t,w,x,\boldsymbol{p})=\mathbb{E}_{t,w,x,\boldsymbol{p}}[\nu(T,W_T^h,X_T,\boldsymbol{\pi}_T)]=\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_t^h],$$

donde se utilizó la condición terminal $\nu(T,\cdot)=\log(\cdot)$. Por lo tanto,

$$\nu(t,w,x,\boldsymbol{p})=\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_t^h]\leq\sup_{h\in\mathcal{A}^{\mathbb{F}}}\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_t^h]=V(t,w,x,\boldsymbol{p}),$$

con lo que concluimos la prueba. ■

Observación 4.2.1. *Respecto a las propiedades de la estrategia óptima, en general comparte las descritas en el caso de información completa, reemplazando los parámetros no observables en* *términos del proceso de filtros, $\boldsymbol{\pi}$.*

4.3. Caso beta neutral

Se estudiará ahora el caso en que las carteras de inversión son neutrales al riesgo de mercado, es decir, estrategias de inversión como se definen a continuación:

Definición 4.3.1. *Una estrategia $h^\beta = (h^{(\beta,m)}, h^{(\beta,1)}, h^{(\beta,2)})$ es beta neutral $\mathbb{F}$-admisible si es autofinanciada, $\mathbb{F}$-predecible y satisface la condición de neutralidad dada en* (3.23) *y la condición de integrabilidad* (3.24)*. Denotaremos por $\mathcal{A}^{\mathbb{F},\beta}$ al conjunto de estrategias beta neutrales $\mathbb{F}$-admisibles.*

Notemos que la dinámica del portafolio asociado a la estrategia h^β está dada por

$$\begin{aligned} dW_t^{h^\beta} &= W_t^{h^\beta}\Big[\big(r + h_t^{(\beta,m)}\mu_m + h_t^{(\beta,1)}\boldsymbol{\mu}_1(X_t)^T\boldsymbol{\pi}_t + h_t^{(\beta,2)}\boldsymbol{\mu}_2(X_t)^T\boldsymbol{\pi}_t\big)dt \\ &\qquad + \sigma_m h_t^{(\beta,m)} dB_t^{(m)} + \big(\sigma_1 h_t^{(\beta,1)} + \rho\sigma_2 h_t^{(\beta,2)}\big)dI_t^{(1)} + \sigma_2\sqrt{1-\rho^2}h_t^{(\beta,2)}dI_t^{(2)}\Big] \\ &= f^\beta(t, X_t, \boldsymbol{\pi}_t)W_t^{h^\beta}dt + \sum_{l=0}^{2} g_l^\beta(t)W_t^{h^\beta}dB_t^l, \end{aligned}$$

con $f^\beta(t, X_t, \boldsymbol{\pi}_t) \coloneqq r + h_t^{(\beta,m)}\mu_m + h_t^{(\beta,1)}\boldsymbol{\mu}_1(X_t)^T\boldsymbol{\pi}_t + h_t^{(\beta,2)}\boldsymbol{\mu}_2(X_t)^T\boldsymbol{\pi}_t$, $(B_0, B_1, B_2) \coloneqq (B^{(m)}, I^{(1)}, I^{(2)})$ y

$$\begin{aligned} g_0^\beta(t) &\coloneqq \sigma_m h_t^{(\beta,m)}, \\ g_1^\beta(t) &\coloneqq \sigma_1 h_t^{(\beta,1)} + \rho\sigma_2 h_t^{(\beta,2)}, \\ g_2^\beta(t) &\coloneqq \sigma_2\sqrt{1-\rho^2}h_t^{(\beta,2)}. \end{aligned}$$

De tal forma que, por el Lema 1.1.1, para $u \le t$, podemos escribir a $W_t^{h^\beta}$ como sigue

$$\begin{aligned} W_t^{h^\beta} &= W_u^{h^\beta}\exp\left(\int_u^t\left[f(s, X_s, \boldsymbol{\pi}_s) - \frac{1}{2}\sum_{l=0}^{2}(g_l^\beta(s))^2\right]ds + \sum_{l=0}^{2}\int_u^t g_l(s)dB_s^l\right) \\ &=: W_u^{h^\beta}\exp(G^\beta(u, t, X_s, \boldsymbol{\pi}_s; u \le s \le t)), \end{aligned} \tag{4.39}$$

donde

$$G^\beta(u, t, X_s, \boldsymbol{\pi}_s; u \le s \le t) \coloneqq \int_u^t\left[f(s, X_s, \boldsymbol{\pi}_s) - \frac{1}{2}\sum_{l=0}^{2}(g_l^\beta(s))^2\right]ds + \sum_{l=0}^{2}\int_u^t g_l(s)dB_s^l.$$

En el siguiente teorema se da la expresión de la estrategia beta neutral óptima y la función de valor del correspondiente problema de control.

Teorema 4.3.1. *Supongamos que* $\lambda_1(y) = \lambda_1$ *y* $\lambda_2(y) = \lambda_2$ *y consideremos un inversionista con función de utilidad logarítmica. Entonces la estrategia beta neutral óptima* $h_*^\beta = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)}) \in \mathcal{A}^{\mathbb{F},\beta}$ *está dada por*

$$\begin{aligned} h_*^{(\beta,1)}(t,x,\boldsymbol{p}) &= -\frac{\lambda_1(x - \boldsymbol{\alpha}_1^T\boldsymbol{p}) + \frac{\beta_1}{\beta_2}\lambda_2(x - \boldsymbol{\alpha}_2^T\boldsymbol{p})}{b_1^2 + \frac{\beta_1^2}{\beta_2^2}b_2^2 + \sigma^2[1 - \frac{\beta_1}{\beta_2}]^2}, \\ h_*^{(\beta,2)}(t,x,\boldsymbol{p}) &= -\frac{\beta_1}{\beta_2}h_*^{(\beta,1)}(t,x,\boldsymbol{p}), \\ h_*^{(\beta,m)}(t,x,\boldsymbol{p}) &= \frac{\mu_m}{\sigma_m^2}. \end{aligned}$$

Además la función de valor es de la forma

$$V(t,w,x,\boldsymbol{p}) = \log(w) + \bar{m}(t)x^2 + \bar{n}(t,\boldsymbol{p})x + \bar{u}(t,\boldsymbol{p}),$$

donde, para $0 \le t \le T$, *la función* $\bar{m}(t)$ *es solución de la ecuación diferencial ordinaria*

$$\bar{m}_t(t) - 2\bar{m}(t)(\lambda_1 + \lambda_2) + \Phi_1 = 0, \tag{4.40}$$

con condición terminal $\bar{m}(T) = 0$. *Y las funciones* $\bar{n}(t,\boldsymbol{p})$ *y* $\bar{u}(t,\boldsymbol{p})$, *para* $0 \le t \le T$, *resuelven el siguiente sistema de ecuaciones diferenciales parciales*

$$\begin{aligned} &\bar{n}_t(t,\boldsymbol{p}) - \bar{n}(t,\boldsymbol{p})(\lambda_1 + \lambda_2) + \sum_{i,j\in\mathcal{E}} \bar{n}_{p_i}(t,\boldsymbol{p})q^{ij}p^j + 2\bar{m}(t)(\Gamma_1 + \lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p} + \lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p}) \\ &\quad + \frac{1}{2}\sum_{i,j\in\mathcal{E}} \bar{n}_{p_ip_j}(t,\boldsymbol{p})(H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})) - \Phi_2(\boldsymbol{p}) = 0, \end{aligned} \tag{4.41}$$

$$\begin{aligned} &\bar{u}_t(t,\boldsymbol{p}) + \Gamma_2\bar{m}(t) + \bar{n}(t,\boldsymbol{p})(\Gamma_1 + \lambda_1\boldsymbol{\alpha}_1^t\boldsymbol{p} + \lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p}) + \Phi_3(\boldsymbol{p}) \\ &\quad + \sum_{i,j\in\mathcal{E}} \bar{u}_{p_i}(t,\boldsymbol{p})q^{ij}p^j + \frac{1}{2}\sum_{i,j\in\mathcal{E}} \bar{u}_{p_ip_j}(t,\boldsymbol{p})(H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})) \\ &\quad + \sum_{i\in\mathcal{E}} \bar{n}_{p_i}(t,\boldsymbol{p})\left[\frac{b_1^2}{\sqrt{\sigma^2 + b_2^2}}H^{i,(1)}(\boldsymbol{p}) - \frac{\sqrt{\sigma^2(b_1^2 + b_2^2) + b_2^2b_1^2}}{\sqrt{\sigma^2 + b_1^2}}H^{i,(2)}(\boldsymbol{p})\right] = 0, \end{aligned} \tag{4.42}$$

con condiciones terminales $\bar{n}(T,\boldsymbol{p}) = 0$ *y* $\bar{u}(T,\boldsymbol{p}) = 0$ *y donde* Γ_1 *y* Γ_2 *están dados como en el Teorema 3.1.1 y*

$$\begin{aligned} \Phi_1 &= \frac{(\beta_2\lambda_1 + \beta_1\lambda_2)^2}{2(\beta_2^2b_1^2 + b_2^2\beta_1^2 + \sigma^2(\beta_1 - \beta_2)^2)}, \\ \Phi_2(\boldsymbol{p}) &= \frac{(\beta_2\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p} + \beta_1\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})(\beta_1\lambda_2 + \beta_2\lambda_1)}{\beta_2^2b_1^2 + b_2^2\beta_1^2 + \sigma^2(\beta_1 - \beta_2)^2}, \\ \Phi_3(\boldsymbol{p}) &= r + \frac{\mu_m^2}{2\sigma_m^2} + \frac{(\beta_2\lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p} + \beta_1\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})^2}{2(\beta_2^2b_1^2 + b_2^2\beta_1^2 + \sigma^2(\beta_1 - \beta_2)^2)}. \end{aligned}$$

En lo anterior $\bar{m}_t, \bar{n}_t$, $\bar{u}_t$, $\bar{n}_{p_i}$ *y* $\bar{u}_{p_i}$ *denotan las derivadas parciales de las funciones* $\bar{m}$, $\bar{n}$ *y* $\bar{u}$, *respectivamente.*

Demostración. Como se ha hecho en los casos anteriores, a partir de la ecuación de Hamilton-Jacobi-Bellman se obtendrá la estrategia *beta neutral* óptima, se propondrá un candidato a la función de valor y finalmente se verificará que éste coincide con la solución a la ecuación HJB.

1. Estrategia Óptima

Supongamos que la función de valor V es de clase $\mathcal{C}^{1,2,2,2}([0,T],\mathbb{R}^+,\mathbb{R},\Delta_K)$, ello implica que se satisface la ecuación de Hamilton-Jacobi-Bellman dada por (3.4), con condición terminal $V(T,\cdot) = \log(w)$. Además, por (4.39), de forma análoga al Teorema 4.2.1, se tiene que $V(t,w,x,\boldsymbol{p}) = \log(w) + v(t,x,\boldsymbol{p})$, para alguna función $v(t,x,\boldsymbol{p})$ con $v(T,\cdot) = 0$.

Así, sustituyendo la función de valor en la expresión del generador dada en (4.19) y por la restricción de neutralidad descrita en (3.23), tenemos que

$$\begin{aligned}
\sup_{h^\beta \in \mathcal{A}^{\mathbb{F},\beta}} \Bigg\{ & v_t(t,x,\boldsymbol{p}) + v_x(t,x,\boldsymbol{p})[\Gamma_1 + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}] + \sum_{i,j\in\mathcal{E}} v_{p_i}(t,x,\boldsymbol{p})q^{ij}p^j \\
& + r + h^{(\beta,m)}\mu_m + h^{(\beta,1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} + h^{(\beta,2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p} + \frac{1}{2}v_{xx}(t,x,\boldsymbol{p})\Gamma_2 \\
& + \frac{1}{2}\sum_{i,j\in\mathcal{E}} v_{p_ip_j}(t,x,\boldsymbol{p})[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})] \\
& - \frac{1}{2}\left[(h^{(\beta,m)}\sigma_m)^2 + (\sigma_1 h^{(\beta,1)} + \rho\sigma_2 h^{(\beta,2)})^2 + \sigma_2^2(1-\rho^2)(h^{(\beta,2)})^2)\right] \\
& + \sum_{i\in\mathcal{E}} v_{xp_i}(t,x,\boldsymbol{p})\left[H^{i,(1)}(\boldsymbol{p})(\sigma_1 - \rho\sigma_2) - H^{i,(2)}(\boldsymbol{p})\sigma_2\sqrt{1-\rho^2})\right] \Bigg\} = 0.
\end{aligned} \tag{4.43}$$

Al contar con la restricción de neutralidad definida en (3.23), para poder obtener el supremo en (4.43), hallaremos primero los puntos críticos mediante el método de multiplicadores de Lagrange y luego se verificará que son máximos con las condiciones de segundo orden.

Entonces, por la condiciones de primer orden, considerando la restricción (3.23) y derivando respecto e $h^{(\beta,m)}$, $h^{(\beta,1)}$ y $h^{(\beta,2)}$, respectivamente de (4.43) se sigue que

$$\frac{\mu_m}{\sigma_m^2} - h_*^{(\beta,m)} = 0, \tag{4.44}$$

$$\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \sigma_1^2 h_*^{(\beta,1)} - \sigma^2 h_*^{(\beta,2)} = \Lambda\beta_1, \tag{4.45}$$

$$\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \sigma_2^2 h_*^{(\beta,2)} - \sigma^2 h_*^{(\beta,1)} = \Lambda\beta_2, \tag{4.46}$$

de donde

$$h_*^{(\beta,m)} = \frac{\mu_m}{\sigma_m^2},$$

y

$$\frac{1}{\beta_1}\big[\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \sigma_1^2 h_*^{(\beta,1)} - \sigma^2 h_*^{(\beta,2)}\big] = \frac{1}{\beta_2}\big[\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \sigma_2^2 h_*^{(\beta,2)} - \sigma^2 h_*^{(\beta,1)}\big],$$

esto es

$$\frac{\boldsymbol{\mu}_1(x)^T\boldsymbol{p}}{\beta_1} - \frac{\boldsymbol{\mu}_2(x)^T\boldsymbol{p}}{\beta_2} = \frac{\sigma^2 h_*^{(\beta,2)}}{\beta_1} - \frac{\sigma_2^2 h_*^{(\beta,2)}}{\beta_2} + \frac{\sigma_1^2 h_*^{(\beta,1)}}{\beta_1} - \frac{\sigma^2 h_*^{(\beta,1)}}{\beta_2}.$$

Utilizando la restricción de neutralidad $h^{(\beta,2)} = -\frac{\beta_1}{\beta_2}h^{(\beta,1)}$, tenemos que

$$\frac{\beta_2\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \beta_1\boldsymbol{\mu}_2(x)^T\boldsymbol{p}}{\beta_1\beta_2} = h_*^{(\beta,1)}\left[\frac{\beta_2^2(\sigma^2+b_1^2) - 2\beta_2\beta_1\sigma^2 + \beta_1^2(\sigma^2+b_2^2)}{\beta_1\beta_2^2}\right],$$

ya que $\sigma_1^2 = \sigma^2 + b_1^2$ y $\sigma_2^2 = \sigma^2 + b_2^2$.

Por lo tanto,

$$\beta_2^2\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \beta_1\beta_2\boldsymbol{\mu}_2(x)^T\boldsymbol{p} = h_*^{(\beta,1)}\big[\sigma^2(\beta_2^2 - 2\beta_1\beta_2 + \beta_1^2) + \beta_2^2 b_1^2 + \beta_1^2 b_2^2\big],$$

de donde concluimos que

$$h_*^{(\beta,1)} = -\frac{\lambda_1(x - \boldsymbol{\alpha}_1^T\boldsymbol{p}) + \frac{\beta_1}{\beta_2}\lambda_2(x - \boldsymbol{\alpha}_2^T\boldsymbol{p})}{b_1^2 + \frac{\beta_1^2}{\beta_2^2}b_2^2 + \sigma^2\big(1 - \frac{\beta_1}{\beta_2}\big)^2},$$

pues $\boldsymbol{\mu}_1(X_t)^T\boldsymbol{\pi}_t = -\lambda_1(X_t - \boldsymbol{\alpha}_1^T\boldsymbol{\pi}_t)$ y $\boldsymbol{\mu}_2(X_t)^T\boldsymbol{\pi}_t = \lambda_2(X_t - \boldsymbol{\alpha}_2^T\boldsymbol{\pi}_t)$.

Además, por (3.23)

$$h_*^{(\beta,2)} = -\frac{\beta_1}{\beta_2}h_*^{(\beta,2)}.$$

Para verificar las condiciones de segundo orden, obteniendo las segundas derivadas obtenemos

$$A^\beta = \begin{pmatrix} -\sigma_m^2 & 0 & 0 \\ 0 & -(\sigma^2+b_1^2) & -\sigma^2 \\ 0 & -\sigma^2 & -(\sigma^2+b_2^2) \end{pmatrix},$$

notemos que A^β es igual a la del Teorema 3.2.1, por lo tanto es definida no positiva y se verifica que $h_* = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)})$ es óptimo.

2. Existencia

Antes de proponer un candidato a la función de valor notemos que, de la ecuación (4.45)

$$\Lambda = \frac{1}{\beta_1}\big[\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \sigma_1^2 h_*^{(\beta,1)} - \sigma^2 h_*^{(\beta,2)}\big],$$

sustituyendo en (4.46)

$$\frac{\beta_2}{\beta_1}\big[\boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \sigma_1^2 h^{(\beta,1)} - \sigma^2 h_*^{(\beta,2)}\big] = \boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \sigma^2 h_*^{(\beta,1)} - \sigma_2^2 h_*^{(\beta,2)},$$

multiplicando por $h_*^{(\beta,2)}$ y utilizando la restricción de neutralidad, obtenemos

$$\begin{aligned}-{h_*}^{(\beta,1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} + \sigma^2(h_*^{(\beta,1)})^2 + b_1^2(h_*^{(\beta,1)})^2 + \sigma^2 h_*^{(\beta,1)}h_*^{(\beta,2)}&\\ - h_*^{(\beta,2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p} + \sigma^2 h_*^{(\beta,2)}h_*^{(\beta,1)} + \sigma^2(h_*^{(\beta,2)})^2 + b_2^2(h_*^{(\beta,2)})^2 = 0.&\end{aligned} \tag{4.47}$$

Por otro lado, sustituyendo $h_* = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)})$ en (4.43) se tiene que

$$\begin{aligned}v_t(t,x,\boldsymbol{p}) &+ v_x(t,x,\boldsymbol{p})[\Gamma_1 + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}] + \sum_{i,j} q^{ij}p^j v_{p_i}(t,x,\boldsymbol{p})\\ &+ r + \frac{\mu_m^2}{2\sigma_m^2} + h_*^{(\beta,1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} + h_*^{(\beta,2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p} - \frac{1}{2}\big[\sigma^2(h_*^{(\beta,1)} + h_*^{(\beta,2)})^2 + (b_1 h_*^{(\beta,1)})^2 + (b_2 h_*^{(\beta,2)})^2\big]\\ &+ \frac{1}{2}v_{xx}(t,x,\boldsymbol{p})\Gamma_2 + \frac{1}{2}\sum_{i,j\in\mathcal{E}} v_{p_ip_j}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big]\\ &+ \sum_{i\in\mathcal{E}} v_{xp_i}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1 - \rho\sigma_2) - \sigma_2\sqrt{1-\rho^2}H^{i,(2)}(\boldsymbol{p})\big] = 0,\end{aligned}$$

y por (4.47)

$$\begin{aligned}v_t(t,x,\boldsymbol{p}) &+ v_x(t,x,\boldsymbol{p})[\Gamma_1 + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}] + \sum_{i,j} q^{ij}p^j v_{p_i}(t,x,\boldsymbol{p})\\ &+ r + \frac{\mu_m^2}{2\sigma_m^2} + \frac{1}{2}\big[h_*^{(\beta,1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} + h_*^{(\beta,2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p}\big] + \frac{1}{2}v_{xx}(t,x,\boldsymbol{p})\Gamma_2\\ &+ \frac{1}{2}\sum_{i,j\in\mathcal{E}} v_{p_ip_j}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big]\\ &+ \sum_{i\in\mathcal{E}} v_{xp_i}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1 - \rho\sigma_2) - \sigma_2\sqrt{1-\rho^2}H^{i,(2)}(\boldsymbol{p})\big] = 0.\end{aligned}$$

Pero, recordemos que $\boldsymbol{\mu}_1(X_t)^T\boldsymbol{\pi}_t = -\lambda_1(X_t - \boldsymbol{\alpha}_1^T\boldsymbol{\pi}_t)$ y $\boldsymbol{\mu}_2(X_t)^T\boldsymbol{\pi}_t = \lambda_2(X_t - \boldsymbol{\alpha}_2^T\boldsymbol{\pi}_t)$, entonces

$$h_*^{(\beta,1)}\boldsymbol{\mu}_1(x)^T\boldsymbol{p} = -\lambda_1(x - \boldsymbol{\alpha}_1^T\boldsymbol{p})\cdot\frac{-\lambda_1(x - \boldsymbol{\alpha}_1^T\boldsymbol{p}) - \frac{\beta_1}{\beta_2}\lambda_2(x - \boldsymbol{\alpha}_2^T\boldsymbol{p})}{b_1^2 + b_2^2\frac{\beta_1^2}{\beta_2^2} + \sigma^2\big(\frac{\beta_2-\beta_1}{\beta_2}\big)^2}$$

$$
\begin{aligned}
&= \frac{\lambda_1^2\beta_2^2(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})(x-\boldsymbol{\alpha}_1^T\boldsymbol{p}) + \beta_1\beta_2\lambda_1\lambda_2(x-\boldsymbol{\alpha}_1^T\boldsymbol{p})(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})}{\beta_2^2 b_1^2 + b_2^2\beta_1^2 + \sigma^2(\beta_2-\beta_1)^2} \\
&= \frac{\beta_2^2\lambda_1^2(x^2 - 2x\boldsymbol{\alpha}_1^T\boldsymbol{p} + (\boldsymbol{\alpha}_1^T\boldsymbol{p})^2) + \beta_1\beta_2\lambda_1\lambda_2(x^2 - x(\boldsymbol{\alpha}_1^T\boldsymbol{p} + \boldsymbol{\alpha}_2^T\boldsymbol{p}) + \boldsymbol{\alpha}_1^T\boldsymbol{p}\boldsymbol{\alpha}_2^T\boldsymbol{p})}{\beta_2^2 b_1^2 + b_2^2\beta_1^2 + \sigma^2(\beta_2-\beta_1)^2},
\end{aligned}
$$

y

$$
\begin{aligned}
h_*^{(\beta,2)}\boldsymbol{\mu}_2(x)^T\boldsymbol{p} &= -\lambda_2(x-\boldsymbol{\alpha}_2^T\boldsymbol{p}) \cdot \frac{\beta_1\big[-\lambda_1(x-\boldsymbol{\alpha}_1^T\boldsymbol{p}) - \frac{\beta_1}{\beta_2}\lambda_2(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})\big]}{\beta_2\big[b_1^2 + b_2^2\frac{\beta_1^2}{\beta_2^2} + \sigma^2\big(\frac{\beta_2-\beta_1}{\beta_2}\big)^2\big]} \\
&= \frac{\beta_1\beta_2\lambda_1\lambda_2(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})(x-\boldsymbol{\alpha}_1^T\boldsymbol{p}) + \beta_1^2\lambda_2^2(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})(x-\boldsymbol{\alpha}_2^T\boldsymbol{p})}{\beta_2^2 b_1^2 + b_2^2\beta_1^2 + \sigma^2(\beta_2-\beta_1)^2} \\
&= \frac{\beta_1\beta_2\lambda_1\lambda_2(x^2 - x(\boldsymbol{\alpha}_1^T\boldsymbol{p} + \boldsymbol{\alpha}_2^T\boldsymbol{p}) + \boldsymbol{\alpha}_1^T\boldsymbol{p}\boldsymbol{\alpha}_2^T\boldsymbol{p}) + \beta_1^2\lambda_2^2(x^2 - 2x\boldsymbol{\alpha}_2^T\boldsymbol{p} + (\boldsymbol{\alpha}_2^T\boldsymbol{p})^2)}{\beta_2^2 b_1^2 + b_2^2\beta_1^2 + \sigma^2(\beta_2-\beta_1)^2},
\end{aligned}
$$

por lo que

$$
\begin{aligned}
&v_t(t,x,\boldsymbol{p}) + v_x(t,x,\boldsymbol{p})[\Gamma_1 + \boldsymbol{\mu}_1(x)^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}] + \sum_{i,j} q^{ij}p^j v_{p_i}(t,x,\boldsymbol{p}) \\
&+ \Phi_1 x^2 - \Phi_2(\boldsymbol{p})x + \Phi_3(\boldsymbol{p}) + \frac{1}{2}v_{xx}(t,x,\boldsymbol{p})\Gamma_2 \\
&+ \frac{1}{2}\sum_{i,j\in\mathcal{E}} v_{p_ip_j}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big] \\
&+ \sum_{i\in\mathcal{E}} v_{xp_i}(t,x,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1 - \rho\sigma_2) - \sigma_2\sqrt{1-\rho^2}H^{i,(2)}(\boldsymbol{p})\big] = 0,
\end{aligned}
\tag{4.48}
$$

con Φ_1, $\Phi_2(\boldsymbol{p})$ y $\Phi_3(\boldsymbol{p})$ definidos como el enunciado del teorema.

Ahora, proponemos como función de valor $v(t,x,\boldsymbol{p}) = \bar{m}(t)x^2 + \bar{n}(t,\boldsymbol{p})x + \bar{u}(t,\boldsymbol{p})$, la cual al sustituirla en (4.48), se satisface

$$
\begin{aligned}
&\bar{m}_t(t)x^2 + \bar{n}_t(t,\boldsymbol{p})x + \bar{u}_t(t,\boldsymbol{p}) + [\Gamma_1 + \boldsymbol{\mu}_1^T\boldsymbol{p} - \boldsymbol{\mu}_2(x)^T\boldsymbol{p}][2\bar{m}(t)x + \bar{n}(t,\boldsymbol{p})] \\
&\quad + \sum_{i,j\in\mathcal{E}} q^{ij}p^j\big[\bar{n}_{p_i}(t,\boldsymbol{p})x + \bar{u}_{p_i}(t,\boldsymbol{p})\big] + \Phi_1 x^2 - \Phi_2(\boldsymbol{p})x + \Phi_3(\boldsymbol{p}) + \Gamma_2\bar{m}(t) \\
&\quad + \frac{1}{2}\sum_{i,j\in\mathcal{E}}\big[\bar{n}_{p_ip_j}(t,\boldsymbol{p})x + \bar{u}_{p_ip_j}(t,\boldsymbol{p})\big]\big[H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p}) + H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p})\big] \\
&\quad + \sum_{i\in\mathcal{E}}\bar{n}_{p_i}(t,\boldsymbol{p})\big[H^{i,(1)}(\boldsymbol{p})(\sigma_1 - \rho\sigma_2) - \sigma_2\sqrt{1-\rho^2}H^{i,(2)}(\boldsymbol{p})\big] = 0,
\end{aligned}
$$

es decir,

$$
\begin{aligned}
&x^2\big[\bar{m}_t(t) - 2\bar{m}(t)(\lambda_1+\lambda_2) + \Phi_1\big] \\
&x\Bigg[\bar{n}_t(t,\boldsymbol{p}) - \bar{n}(t,\boldsymbol{p})(\lambda_1+\lambda_2) + \sum_{i,j\in\mathcal{E}}\bar{n}_{p_i}(t,\boldsymbol{p})q^{ij}p^j + 2\bar{m}(t)(\Gamma_1 + \lambda_1\boldsymbol{\alpha}_1^T\boldsymbol{p} + \lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})
\end{aligned}
$$

$$+\frac{1}{2}\sum_{i,j\in\mathcal{E}}\bar{n}_{p_ip_j}(t,\boldsymbol{p})(H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p})+H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p}))-\Phi_2(\boldsymbol{p})\Big]$$
$$+\bar{u}_t(t,\boldsymbol{p})+\Gamma_2\bar{m}(t)+\bar{n}(t,\boldsymbol{p})(\Gamma_1+\lambda_1\boldsymbol{\alpha}_1^t\boldsymbol{p}+\lambda_2\boldsymbol{\alpha}_2^T\boldsymbol{p})+\Phi_3(\boldsymbol{p})$$
$$+\sum_{i,j\in\mathcal{E}}\bar{u}_{p_i}(t,\boldsymbol{p})q^{ij}p^j+\frac{1}{2}\sum_{i,j\in\mathcal{E}}\bar{u}_{p_ip_j}(t,\boldsymbol{p})(H^{i,(1)}(\boldsymbol{p})H^{j,(1)}(\boldsymbol{p})+H^{i,(2)}(\boldsymbol{p})H^{j,(2)}(\boldsymbol{p}))$$
$$+\sum_{i\in\mathcal{E}}\bar{n}_{p_i}(t,\boldsymbol{p})\Big[\frac{b_1^2}{\sqrt{\sigma^2+b_2^2}}H^{i,(1)}(\boldsymbol{p})-\frac{\sqrt{\sigma^2(b_1^2+b_2^2)+b_2^2b_1^2}}{\sqrt{\sigma^2+b_1^2}}H^{i,(2)}(\boldsymbol{p})\Big]=0.$$

Por lo que, si las funciones $\bar{m}(t)$, $\bar{n}(t,\boldsymbol{p})$ y $\bar{u}(t,\boldsymbol{p})$ resuelven los sistemas de ecuaciones dados en (4.40), (4.41) y (4.42), respectivamente, con condiciones terminales $\bar{m}(T)=0$, $\bar{n}(T,\boldsymbol{p})=0$ y $\bar{u}(T,\boldsymbol{p})=0$, entonces la función $v(t,x,\boldsymbol{p})=\bar{m}(t)x^2+\bar{n}(t,\boldsymbol{p})+\bar{u}(t,\boldsymbol{p})$ satisface la ecuación HJB. Por lo tanto, la función de valor es de la forma $V(t,w,x,\boldsymbol{p})=\log(w)+\bar{m}(t)x^2+\bar{n}(t,\boldsymbol{p})x+\bar{u}(t,\boldsymbol{p})$.

3. Verificación

Sea $\nu(t,w,x,\boldsymbol{p})$ de clase $C^{1,2,2,2}([0,T],\mathbb{R}^+,\mathbb{R},\Delta_K)$ y con crecimiento cuadrático una solución a la ecuación HJB, dada en (4.43), y $h^\beta\in\mathcal{A}^{\mathbb{F},\beta}$ una estrategia admisible.

Aplicando la fórmula de Itô a $\nu(T,W_T,X_T,\boldsymbol{\pi}_T)$, de las expresiones dadas en (4.36) y (4.37), y utilizando la hipótesis de neutralidad, tenemos que

$$\begin{aligned}\nu(T,W_T^h,X_T,\boldsymbol{\pi}_T)&=\nu(t,w,x,\boldsymbol{p})+\int_t^T\mathcal{L}_{\mathbb{F}}^{h^\beta}\nu(s,W_s^h,X_s,\boldsymbol{\pi}_s)ds\\&+\int_t^T\big[W_s^h\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)h_s^{(m)}+\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\beta_1-\beta_2)\big]\sigma_m dB_s^{(m)}\\&+\int_t^T\big[W_s^h\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\sigma_1h_s^{(1)}+\rho\sigma_2h_s^{(2)})+\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\sigma_1-\rho\sigma_2)\big]dI_s^{(1)}\\&+\int_t^T\sum_{i\in\mathcal{E}}\nu_{p_i}(s,W_s^h,X_s,\boldsymbol{\pi}_s)H^{i,(1)}(\boldsymbol{\pi}_s)dI_s^{(1)}\\&+\int_t^T\sigma_2\sqrt{1-\rho^2}\big[W_s^hh_s^{(2)}\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)-\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)\big]dI_s^{(2)}\\&+\int_t^T\sum_{i\in\mathcal{E}}\nu_{p_i}(s,W_s^h,X_s,\boldsymbol{\pi}_s)H^{i,(2)}(\boldsymbol{\pi}_s)dI_s^{(2)},\end{aligned}\tag{4.49}$$

donde $\mathcal{L}_{\mathbb{F}}^{h^\beta}$ denota el generador del proceso $(W,X,\boldsymbol{\pi})$ definido en (4.19).

Además, como ν satisface la ecuación HJB, entonces

$$\begin{aligned}\nu(T,W_T^h,X_T,\boldsymbol{\pi}_T)&\le\nu(t,w,x,\boldsymbol{p})+\int_t^T\sum_{i\in\mathcal{E}}\nu_{p_i}(s,W_s^h,X_s,\boldsymbol{\pi}_s)H^{i,(1)}(\boldsymbol{\pi}_s)dI_s^{(1)}\\&+\int_t^T\big[W_s^h\nu_w(s,W_s^h,X_s,\boldsymbol{\pi}_s)h_s^{(m)}+\nu_x(s,W_s^h,X_s,\boldsymbol{\pi}_s)(\beta_1-\beta_2)\big]\sigma_m dB_s^{(m)}\end{aligned}$$

$$+\int_t^T \left[W_s^h \nu_w(s, W_s^h, X_s, \boldsymbol{\pi}_s)(\sigma_1 h_s^{(1)} + \rho\sigma_2 h_s^{(2)}) + \nu_x(s, W_s^h, X_s, \boldsymbol{\pi}_s)(\sigma_1 - \rho\sigma_2)\right] dI_s^{(1)}$$
$$+\int_t^T \sigma_2\sqrt{1-\rho^2}\left[W_s^h h_s^{(2)} \nu_w(s, W_s^h, X_s, \boldsymbol{\pi}_s) - \nu_x(s, W_s^h, X_s, \boldsymbol{\pi}_s)\right] dI_s^{(2)}$$
$$+\int_t^T \sum_{i\in\mathcal{E}} \nu_{p_i}(s, W_s^h, X_s, \boldsymbol{\pi}_s) H^{i,(2)}(\boldsymbol{\pi}_s) dI_s^{(2)}.$$

Si localizamos como se hizo en el Teorema 4.2.1 se puede ver que las integrales respecto a $B^{(m)}$, $I^{(1)}$ y $I^{(2)}$ tienen esperanza cero. Por lo que, de la expresión anterior, $\mathbb{E}_{t,w,x,\boldsymbol{p}}[\nu(T, W_T^{h^\beta}, X_t, \boldsymbol{\pi}_T)] \leq \nu(t, w, x, \boldsymbol{p})$ y por la condición terminal $\nu(T, \cdot) = \log(w)$ se sigue que $\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_t^{h^\beta}] \leq \nu(t, w, x, \boldsymbol{p})$. Luego, tomando supremo sobre las estrategias admisibles h^β, $V(t, w, x, \boldsymbol{p}) \leq \nu(t, w, x, \boldsymbol{p})$.

Por otro lado, si $h_* = (h_*^{(\beta,m)}, h_*^{(\beta,1)}, h_*^{(\beta,2)})$ es la estrategia *beta neutral* óptima, de (4.49), análogo al Teorema 4.2.1, obtenemos

$$\mathbb{E}_{t,w,x,\boldsymbol{p}}[\nu(T, W_T^{h^\beta}, X_t, \boldsymbol{\pi}_T)] = \nu(t, w, x, \boldsymbol{p}),$$

y por la condición terminal,

$$\mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_T^h] = \nu(t, w, x, \boldsymbol{p}),$$

por lo que

$$V(t, w, x, \boldsymbol{p}) = \sup_{h^\beta \in \mathcal{A}^{\mathbb{F},\beta}} \mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_T^h] \geq \mathbb{E}_{t,w,x,\boldsymbol{p}}[\log W_T^h] = \nu(t, w, x, \boldsymbol{p}),$$

de donde se sigue el resultado. ■

Observaciones respecto a la estrategia óptima

1. Al igual que en el caso de información completa, notemos que el cociente $\frac{\beta_1}{\beta_2}$ tiene la misma interpretación que ϱ_2 en el Teorema 3.1.1.

2. Si consideramos $\beta_1 = \beta_2$ volvemos a obtener una estrategia *delta neutral*, es decir,

$$h_*^{(1)}(t, x, \boldsymbol{p}) = -h_*^{(2)}(t, x, \boldsymbol{p}) = -\frac{\lambda_1(x - \boldsymbol{\alpha}_1^T \boldsymbol{p}) + \lambda_2(x - \boldsymbol{\alpha}_2^T \boldsymbol{p})}{b_1^2 + b_2^2},$$
$$h_*^{(m)}(t, x, \boldsymbol{p}) = \frac{\mu_m}{\sigma_m^2}.$$

Conclusiones

Este trabajo resulta ser una monografía del artículo *Optimal convergence trading with unobservable pricing errors* [2], donde se estudia el problema de optimización de portafolios con estrategias convergentes. El modelo propuesto es una modificación al expuesto en [8], la cual supone que los *errores* en los activos cointegrados están regidos por una cadena de Markov. De aquí surgen dos enfoques: información parcial, donde se supone que la cadena no es observable por el inversionista; e información completa, bajo el supuesto de que tal cadena sí es observable. Además, se estudia el caso en que las carteras de inversión se restringen a ser *beta neutrales.*

La herramienta principal para resolver el problema de optimización fue control estocástico, particularmente la técnica descrita en la Sección 1.2. Los resultados obtenidos fueron:

1. Caso información completa: se obtuvo la estrategia óptima del problema de control, tanto para el caso clásico como para el caso *beta neutral.* Asimismo, se caracteriza la función de valor como la única solución a un sistema de ecuaciones diferenciales ordinarias derivado de la ecuación de Hamilton-Jacobi-Bellman.

2. Caso información parcial: para este enfoque, al ser la cadena de Markov no observable por el inversionista, primero se pasó de un problema de información parcial a uno equivalente de información completa. Lo anterior vía el *problema de filtros*, el cual permitió obtener la probabilidad condicional de la cadena de Markov dada la información a la cual tiene acceso el inversionista. Una vez se planteó el problema equivalente bajo información completa, análogamente al caso anterior, se encontró la estrategia óptima y se caracterizó la función de valor como la única solución a un sistema de ecuaciones diferenciales parciales derivado de la ecuación de Hamilton-Jacobi-Bellman.

Aunado a lo anterior, al ser este un trabajo monográfico, durante el desarrollo me fue posible profundizar conocimientos en ciertos campos, particularmente el problema de optimización de portafolios desde un punto de vista de la teoría de control estocástico, así como ciertos conceptos referentes a cálculo estocástico.

Bibliografía

[1] Allinger, D. F. y S. K. Mitter. (1981). New results on the innovations problem for non-linear filtering. *Stochastics: An International Journal of Probability and Stochastic Processes*, 4(4):339–348.

[2] Altay, S., K. Colaneri y Z. Eksi. (2020). Optimal convergence trading with unobservable pricing errors. Ann Oper Res. *https://doi.org/10.1007/s10479-020-03647-z*

[3] Bain, A. y D. Crisan. (2009). Fundamentals of Stochastic Filtering. *Springer.*

[4] Evans, L. (2013). An Introduction to Stochastic Differential Equations. *American Mathematical Society.*

[5] Ikeda, N. y S. Watanabe. (1981). Stochastic Differential Equations and Diffusion Processes. *Kodanasha Ltd.*

[6] Jacod, J. y A. N. Shiryaev. (1987). Limit Theorems for Stochastic Processes, volume 2003. *Springer.*

[7] Korn, R. (1997). Optimal Portfolios: Stochastic Models for Optimal Investment and Risk Management in Continuous Time. *World Scientific*

[8] Liu, J. y A. Timmermann. (2013). Optimal convergence trade strategies. *The Review of Financial Studies*, 26(4):1048–1086.

[9] Le Gall, J. F. (2016). Brownian Motion, Martingales, and Stochastic Calculus. *Springer International Publishing*

[10] Merton, R. (1971). Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory*, 3:373413.

[11] Levy, H. (2012). The Capital Asset Pricing Model in the 21st Century. Analytical, Empirical, and Behavioral Perspectives. *Cambridge University Press.*

[12] Pham, H. (2009). Continuous-time Stochastic Control and Optimization with Financial Applications. *Springer Publishing Company.*

[13] Protter, P. (1990). Stochastic Integration and Differential Equations. A New Approach. *Springer-Verlag.*

[14] Shreve, S. E. (2004). Stochastic Calculus for Finance II: Continuous-time Models (Vol. 11). *Springer Science & Business Media.*

[15] Touzi, N. (2013). Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE. *New York, Springer.*