



UNIVERSIDAD DE BUENOS AIRES  
Facultad de Ciencias Exactas y Naturales  
Departamento de Matemática

Tesis de Licenciatura

Juegos de campo medio

Lucía María Chiappara

**Director:** Juan Pablo Pinasco

Marzo de 2011



# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Control óptimo determinístico</b>	<b>3</b>
2.1. El problema . . . . .	3
2.2. Programación Dinámica . . . . .	4
2.2.1. Principio de Programación Dinámica . . . . .	4
2.2.2. Ecuación de Hamilton-Jacobi-Bellman . . . . .	7
2.2.3. Diseño del Control Óptimo . . . . .	9
2.3. Soluciones Viscosas . . . . .	10
2.3.1. Consistencia . . . . .	11
2.3.2. Unicidad . . . . .	13
2.3.3. Fórmula de Lax-Hopf . . . . .	14
<b>3. Control óptimo estocástico</b>	<b>17</b>
3.1. El problema . . . . .	17
3.2. Ecuaciones y cálculo estocástico . . . . .	18
3.2.1. Regla de la Cadena de Itô . . . . .	20
3.3. Programación Dinámica . . . . .	23
3.3.1. Ecuación de Hamilton-Jacobi-Bellman . . . . .	23
3.3.2. Diseño del Control Óptimo . . . . .	26
<b>4. Juegos Diferenciales</b>	<b>27</b>
4.1. Definiciones . . . . .	27
4.2. Propiedades del Valor Superior e Inferior . . . . .	30
<b>5. Juegos de campo medio</b>	<b>33</b>
5.1. ¿Cuándo empieza una reunión? . . . . .	33
5.2. El problema de la calefacción . . . . .	38
5.2.1. El modelo . . . . .	38
5.2.2. Distribución de los jugadores . . . . .	38
5.2.3. Costos . . . . .	40

5.2.4.	El problema . . . . .	41
5.2.5.	Existencia de un ínfimo . . . . .	42
<b>Bibliografía</b>		<b>47</b>

# Capítulo 1

## Introducción

Los juegos de *campo medio* (Mean-field games) son una rama reciente de la Teoría de Juegos, introducida por Jean Michel Lasry y Pierre-Louis Lions en [10, 11, 12]. En un juego diferencial de  $N$  jugadores, cada agente intenta, individualmente, optimizar cierto funcional que va a depender de su posición (espacial y temporal), pero se verá también influido por las decisiones de los demás. Usualmente, esta situación se representa con un sistema de ecuaciones de Hamilton-Jacobi-Bellman-Isaacs, y como veremos en el Capítulo §4, se torna intratable con más de dos jugadores.

Por otra parte, si consideramos un número de jugadores  $N$  muy grande, las cosas se simplifican. Considerando aquellos juegos donde el rol de los jugadores es intercambiable, a medida que la cantidad de jugadores crece, el efecto de las decisiones individuales se va tornando insignificante sobre la masa total. Si tomamos el límite cuando  $N$  tiende infinito, en lugar de la influencia puntual de cada uno, tendremos un efecto que dependerá de la *densidad* local de jugadores. Ahora, cada jugador tomará sus decisiones basándose en su situación y en la de sus vecinos, quienes a su vez decidirán de la misma forma. De aquí proviene el nombre de estos juegos, se introduce un *campo medio* para describir la interacción entre los agentes: el comportamiento de los agentes individuales genera un campo, que, a su vez, tiene un efecto sobre cada uno de ellos.

Esta idea fue originalmente usada en Física, para modelar sistemas de muchas partículas. La diferencia principal, y el desafío, de trasladar este concepto a juegos es que aquí las partículas son agentes racionales que tienen la capacidad de interactuar estratégicamente: cada agente desarrolla estrategias observando a sus vecinos, y viceversa. Esto se ve reflejada en un funcional a optimizar, que genera una ecuación de tipo Hamilton-Jacobi, para describir el cálculo de la trayectoria óptima de cada agente.

Bajo ciertas hipótesis, entonces, podemos representar un juego de campo medio con un sistema de dos ecuaciones en derivadas parciales: una ecuación de Kolmogorov (forward), que describirá como evoluciona la densidad de jugadores a medida que avanza el tiempo y éstos se reubican, y una de Hamilton-Jacobi. Esta última evolucionará hacia atrás en el tiempo, y se desprende del concepto de programación dinámica.

En esencia, no existe aún una teoría de juegos de campo medio, sino un conjunto de conceptos, herramientas matemáticas, simulaciones y algoritmos (no siempre determinísticos) que intentan modelar situaciones donde los agentes toman decisiones en un contexto de interacción estratégica. Estos modelos se aplican en diversas áreas, como macro y microeconomía, sociología, ingeniería o arquitectura. Los ejemplos que comúnmente se eligen, podrían ser calificados como “modelos de juguete”, en el sentido de que no se espera que se apliquen literalmente en la vida real, sino que su estructura sea un arquetipo de numerosas situaciones que sí se presentan en la realidad.

En este trabajo, para estudiarlos, veremos primero algunas de las herramientas necesarias. En el Capítulo §2 consideramos un modelo con un solo jugador, que elige su trayectoria en un intervalo de tiempo, intentando maximizar cierto pago. Este es un problema clásico de control óptimo. Usando el *método de Programación Dinámica* convertiremos el problema en otro, consistente en resolver una ecuación en derivadas parciales.

Como dicha ecuación no tiene por qué tener solución en un sentido clásico, definiremos entonces un concepto de solución más débil, el concepto de *solución viscosa*. Veremos que el problema planteado tiene solución única, al menos en este sentido. Para este capítulo nos basamos principalmente en los libros de L. C. Evans [3] y los capítulos 3 y 10 de [2].

En el Capítulo §3 consideraremos el modelo anterior, pero agregando ruido a la trayectoria del jugador. Esto resultará en un problema de control óptimo estocástico y adaptaremos lo hecho antes a esta situación. Un detalle interesante que vale la pena destacar es que la ecuación que se obtiene ahora es una ecuación parabólica. Seguimos aquí el libro de Evans de control [3], y sus notas sobre ecuaciones diferenciales estocásticas [4].

En el Capítulo §4 definiremos juegos diferenciales para dos jugadores (la extensión a un número  $N$  de jugadores es inmediata). Veremos aquí que la solución se reduce a un problema de control para dos jugadores.

Cuando los jugadores disponen de finitas estrategias, y el juego es de suma cero, el Teorema Minimax de Von Neumann asegura que las estrategias minimax y maximin de cada jugador dan el mismo valor para el juego. Para juegos diferenciales no es cierto, y se obtienen sub- y supersoluciones viscosas que, bajo ciertas hipótesis en el funcional a optimizar, coinciden. Estos resultados corresponden al trabajo de L. C. Evans y P. E. Souganidis [5].

Finalmente, en el capítulo §5 presentaremos distintos ejemplos, debidos principalmente a Aimé Lachapelle [9] y Oliver Gueant [7, 8]. Uno de los ejemplos que veremos en este trabajo es “el problema de la calefacción”, y si bien la situación puede parecer artificial, es interesante conceptualmente, pues combina varios conceptos económicos. Veremos también un problema de optimización para el tiempo de llegada a una reunión que comenzará al obtenerse el quorum fijado en un cierto porcentaje de los agentes.

# Capítulo 2

## Control óptimo determinístico

### 2.1. El problema

Dados un conjunto  $A \subseteq \mathbb{R}^m$ , una función  $f : \mathbb{R}^n \times A \rightarrow \mathbb{R}^n$  continua Lipchitz, y una condición inicial  $x_0 \in \mathbb{R}^n$ , queremos resolver el siguiente sistema dinámico

$$\begin{cases} \dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \alpha(t)) \\ \mathbf{x}(0) = x_0 \end{cases} \quad (2.1.1)$$

para todo  $t > 0$ , donde  $\alpha(\cdot)$  es una función que consideraremos fija. Llamaremos *respuesta del sistema* a la trayectoria solución  $\mathbf{x}(\cdot)$ . Observemos que  $\mathbf{x}(\cdot) = \mathbf{x}(\cdot, \alpha(\cdot), x_0)$ , es decir, la respuesta del sistema cambia según la elección de  $\alpha(\cdot)$  ó de la condición inicial. Diremos que  $\alpha(\cdot)$  es un *control* para (2.1.1).

**Definición 2.1.1.** *Definimos la colección de controles admisibles como*

$$\mathcal{A} = \{\alpha : [0, \infty) \rightarrow A \mid \alpha(\cdot) \text{ medible}\}.$$

El objetivo será determinar cuál es el *mejor* control para el sistema. Para ello, necesitamos especificar algún criterio que nos permita compararlos.

**Definición 2.1.2.** *Dados  $T > 0$ , el tiempo terminal,  $r : \mathbb{R}^n \times A \rightarrow \mathbb{R}$  y  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  dos funciones que llamaremos el pago variable y el pago terminal, respectivamente, definimos el funcional de pago como*

$$P[\alpha(\cdot)] := \int_0^T r(\mathbf{x}(t), \alpha(t)) dt + g(\mathbf{x}(T)) \quad (2.1.2)$$

donde  $\mathbf{x}(\cdot)$  es la solución de (2.1.1) para el control  $\alpha(\cdot)$ .

Luego, el problema básico será encontrar  $\alpha^*(\cdot)$  que maximice el pago. Es decir,

$$P(\alpha^*(\cdot)) \geq P(\alpha(\cdot)), \quad \forall \alpha(\cdot) \in \mathcal{A}. \quad (2.1.3)$$

A dicho control lo llamaremos *óptimo*.

## 2.2. Programación Dinámica

El método de programación dinámica provee una forma de diseñar un control óptimo, utilizando la función valor. Dicha función debe satisfacer una ecuación en derivadas parciales, que derivaremos en esta sección. Luego, habremos transformado nuestro problema de optimización en otro, resolver una ecuación en derivadas parciales: Si somos capaces de encontrar una solución de esta ecuación, el método de programación dinámica nos da una receta para obtener un control óptimo para el problema planteado.

Consideremos una familia más grande de problemas, haciendo variar el intervalo de tiempo donde resolvemos el problema, y la posición inicial:

$$\begin{cases} \dot{\mathbf{x}}(s) &= f(\mathbf{x}(s), \alpha(s)) & (t \leq s \leq T) \\ \mathbf{x}(t) &= x \end{cases} \quad (2.2.1)$$

con

$$P_{x,t}[\alpha(\cdot)] = \int_t^T r(\mathbf{x}(s), \alpha(s)) dt + g(\mathbf{x}(T)). \quad (2.2.2)$$

Estudiamos el problema de arriba para todas las posibles elecciones del tiempo inicial  $0 \leq t \leq T$ , y todas las condiciones iniciales  $x \in \mathbb{R}^n$ .

**Definición 2.2.1.** Para  $x \in \mathbb{R}^n$ ,  $0 \leq t \leq T$ , se define la función valor  $v(x, t)$  como el mayor pago posible que se puede obtener si se empieza a tiempo  $t$  en  $x \in \mathbb{R}^n$ . Es decir,

$$v(x, t) := \sup_{\alpha \in \mathcal{A}} \{P_{x,t}[\alpha(\cdot)]\}, \quad x \in \mathbb{R}^n, 0 \leq t \leq T. \quad (2.2.3)$$

Observemos que

$$v(x, T) = g(x) \quad (x \in \mathbb{R}^n). \quad (2.2.4)$$

En lo que sigue, veremos qué ecuación satisface la función  $v$ .

### 2.2.1. Principio de Programación Dinámica

El primer resultado dará una condición de optimalidad:

**Teorema 2.2.1** (Principio de Programación Dinámica). Para cada  $h > 0$ , suficientemente chico tal que  $t + h \leq T$ , se tiene

$$v(x, t) = \sup_{\alpha(\cdot) \in \mathcal{A}} \int_t^{t+h} r(\mathbf{x}(s), \alpha(s)) ds + v(\mathbf{x}(t+h), t+h) \quad (2.2.5)$$

donde  $\mathbf{x}(\cdot)$  es la solución de (2.2.1) para el control  $\alpha(\cdot)$ .

*Demostración.* Vamos a fijar un  $\varepsilon > 0$  arbitrario, y probemos primero que

$$v(x, t) \geq \sup_{\alpha(\cdot) \in \mathcal{A}} \int_t^{t+h} r(\mathbf{x}(s), \alpha(s)) ds + v(\mathbf{x}(t+h), t+h) - \varepsilon.$$

Elegimos  $\alpha_1(\cdot) \in \mathcal{A}$  y resolvemos (2.2.1) para los tiempos  $t < s < t+h$ . Esto nos da una trayectoria  $\mathbf{x}_1(\cdot)$ . Ahora, partiendo del punto  $\mathbf{x}_1(t+h)$  y al tiempo  $t+h$ , supongamos que tenemos un control óptimo. Entonces tenemos definida  $v(\mathbf{x}_1(t+h), t+h)$ . Como  $v$  es un supremo, dado  $\varepsilon < 0$ , existe  $\alpha_2(\cdot) \in \mathcal{A}$  tal que

$$v(\mathbf{x}_1(t+h), t+h) - \varepsilon \leq \int_{t+h}^T r(\mathbf{x}_2(s), \alpha_2(s)) ds + g(x_2(T))$$

donde  $\mathbf{x}_2(\cdot)$  es la solución de

$$\begin{cases} \dot{\mathbf{x}}_2(s) = f(\mathbf{x}_2(s), \alpha_2(s)) & t+h \leq s \leq T \\ \mathbf{x}_2(t+h) = \mathbf{x}_1(t+h) \end{cases} \quad (2.2.6)$$

Definimos el control

$$\alpha_3(s) := \begin{cases} \alpha_1(s) & \text{si } s \in [t, t+h) \\ \alpha_2(s) & \text{si } s \in [t+h, T] \end{cases}$$

y resolvemos (2.2.1) para  $\alpha_3(\cdot)$ , en el intervalo de tiempo  $(t, T)$ , con la condición inicial  $\mathbf{x}_3(t) = x$ . Por la unicidad de la solución de la ecuación ordinaria (2.2.1), tenemos:

$$\mathbf{x}_3(s) = \begin{cases} \mathbf{x}_1(s) & \text{si } s \in [t, t+h) \\ \mathbf{x}_2(s) & \text{si } s \in [t+h, T] \end{cases}$$

Entonces, por la definición de  $v$ ,

$$\begin{aligned} v(x, t) &\geq P_{x,t}[\alpha_3(\cdot)] \\ &= \int_t^T r(\mathbf{x}_3(s), \alpha_3(s)) ds + g(\mathbf{x}(T)) \\ &= \int_t^{t+h} r(\mathbf{x}_1(s), \alpha_1(s)) ds + \int_{t+h}^T r(\mathbf{x}_2(s), \alpha_2(s)) ds + g(\mathbf{x}_2(T)) \\ &\geq \int_t^{t+h} r(\mathbf{x}_1(s), \alpha_1(s)) ds + v(\mathbf{x}_1(t+h), t+h) - \varepsilon, \end{aligned}$$

donde la última desigualdad vale por la definición de  $\alpha_2$ .

Como  $\alpha_1(\cdot) \in \mathcal{A}$  fue elegido arbitrariamente, tenemos

$$v(x, t) \geq \sup_{\alpha(\cdot) \in \mathcal{A}} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \alpha(s)) ds + v(\mathbf{x}(t+h), t+h) \right\} - \varepsilon,$$

y esto vale para todo  $\varepsilon > 0$ . Tomamos  $\varepsilon \rightarrow 0$ , y resulta

$$v(x, t) \geq \sup_{\alpha(\cdot) \in \mathcal{A}} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \alpha(s)) ds + v(\mathbf{x}(t+h), t+h) \right\}. \quad (2.2.7)$$

Para demostrar la otra desigualdad, tomamos de nuevo  $\varepsilon > 0$ , y elegimos  $\alpha_4(\cdot) \in \mathcal{A}$  tal que

$$v(x, t) - \varepsilon \leq \int_t^T r(\mathbf{x}_4(s), \alpha_4(s)) ds + g(\mathbf{x}_4(T)),$$

con  $\mathbf{x}_4(\cdot)$  la solución de la ecuación (2.2.1) en el intervalo  $[t, T]$ , con el control  $\alpha_4(\cdot)$ , y la condición inicial  $\mathbf{x}_4(t) = x$ .

Como  $v(\mathbf{x}(t+h), t+h)$  es un supremo,

$$v(\mathbf{x}_4(t+h), t+h) \geq \int_{t+h}^T r(\mathbf{x}_4(s), \alpha_4(s)) ds + g(\mathbf{x}_4(T)).$$

Juntando las dos desigualdades anteriores se tiene

$$\begin{aligned} v(x, t) - \varepsilon &\leq \int_t^T r(\mathbf{x}_4(s), \alpha_4(s)) ds + g(\mathbf{x}_4(T)) \\ &= \int_t^{t+h} r(\mathbf{x}_4(s), \alpha_4(s)) ds + \int_{t+h}^T r(\mathbf{x}_4(s), \alpha_4(s)) ds + g(\mathbf{x}_4(T)) \\ &\leq \int_t^{t+h} r(\mathbf{x}_4(s), \alpha_4(s)) ds + v(\mathbf{x}_4(t+h), t+h). \end{aligned}$$

Tomando supremo en el lado derecho de la igualdad,

$$v(x, t) - \varepsilon \leq \sup_{\alpha(\cdot) \in \mathcal{A}} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \alpha(s)) ds + v(\mathbf{x}(t+h), t+h) \right\}.$$

Como vale para todo  $\varepsilon > 0$  arbitrario, tomando  $\varepsilon \rightarrow 0$ ,

$$v(x, t) \leq \sup_{\alpha(\cdot) \in \mathcal{A}} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \alpha(s)) ds + v(\mathbf{x}(t+h), t+h) \right\}. \quad (2.2.8)$$

Por último, por las desigualdades (2.2.7) y (2.2.8),

$$v(x, t) = \sup_{\alpha(\cdot) \in \mathcal{A}} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \alpha(s)) ds + v(\mathbf{x}(t+h), t+h) \right\},$$

lo que finaliza la demostración. □

### 2.2.2. Ecuación de Hamilton-Jacobi-Bellman

Ahora queremos deducir una ecuación diferencial para  $v$ . Para esto, estudiaremos la versión infinitesimal del Principio de Programación Dinámica.

**Teorema 2.2.2.** *Supongamos  $v \in C^1(\mathbb{R}^n \times (0, T))$ . Entonces,  $v$  es solución de la ecuación de Hamilton-Jacobi-Bellman (HJB)*

$$v_t(x, t) + \max_{a \in A} \{f(x, a) \cdot \nabla_x v(x, t) + r(x, a)\} = 0, \quad (2.2.9)$$

en  $\mathbb{R}^n \times (0, T)$ , con la condición terminal

$$v(x, T) = g(x) \quad x \in \mathbb{R}^n.$$

*Demostración.* Elegimos  $\alpha(s) \equiv a$  un control constante, con  $a \in A$ . Por el Principio de Programación Dinámica tenemos

$$v(x, t) \geq \int_t^{t+h} r(\mathbf{x}(s), a) ds + v(\mathbf{x}(t+h), t+h),$$

para cualquier  $h > 0$ .

Si pasamos restando el término de la izquierda y dividimos por  $h$ , nos queda

$$\frac{v(\mathbf{x}(t+h), t+h) - v(x, t)}{h} + \frac{1}{h} \int_t^{t+h} r(\mathbf{x}(s), a) ds \leq 0.$$

Queremos ver que, cuando tomamos  $h \rightarrow 0$ , resulta

$$v_t(x, t) + \nabla_x v(x, t) \cdot \dot{\mathbf{x}}(t) + r(x, a) \leq 0.$$

Hagamos esta cuenta con cuidado. Primero acotaremos el cociente incremental, y luego aplicaremos el Teorema Fundamental del Cálculo a la integral.

Recordemos que  $\mathbf{x}(s) = (x_1(s), \dots, x_n(s))$ . Además, por ser solución de la ecuación (2.2.1), es continua en todas sus coordenadas, y  $\mathbf{x}(t) = x$ . Para mostrar la idea, y por simplicidad, alcanza con ver el caso  $n = 2$ :

$$\begin{aligned} \frac{v(\mathbf{x}(t+h), t+h) - v(x, t)}{h} &= \frac{v(\mathbf{x}(t+h), t+h) - v(\mathbf{x}(t+h), t)}{h} \\ &\quad + \frac{v(\mathbf{x}(t+h), t) - v(x, t)}{h}, \end{aligned}$$

entonces, el primer sumando resulta

$$v_t(\mathbf{x}(t), t) = v_t(x, t).$$

Para el segundo,

$$\begin{aligned} \frac{v(\mathbf{x}(t+h), t) - v(x, t)}{h} &= \frac{v((x_1(t+h), x_2(t+h)), t) - v((x_1(t), x_2(t)), t)}{h} \\ &= \frac{v((x_1(t+h), x_2(t+h)), t) - v(x_1(t), x_2(t+h), t)}{h} \\ &\quad + \frac{v((x_1(t), x_2(t+h)), t) - v((x_1(t), x_2(t)), t)}{h} \end{aligned}$$

Si  $x_1(t+h) - x_1(t) \neq 0$ , podemos reescribir esta expresión como

$$\begin{aligned} &\frac{v((x_1(t+h), x_2(t+h)), t) - v(x_1(t), x_2(t+h), t)}{x_1(t+h) - x_1(t)} \frac{x_1(t+h) - x_1(t)}{h} + \\ &+ \frac{v((x_1(t), x_2(t+h)), t) - v((x_1(t), x_2(t)), t)}{x_2(t+h) - x_2(t)} \frac{x_2(t+h) - x_2(t)}{h}, \end{aligned}$$

y tomando límite  $h \rightarrow 0$ , queda

$$v_{x_1}(x, t)\dot{x}_1(t) + v_{x_2}(x, t)\dot{x}_2(t) = \nabla_x v(x, t) \cdot \dot{\mathbf{x}}(t).$$

Por otro lado,

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_t^{t+h} r(\mathbf{x}(s), a) ds = r(\mathbf{x}(t), a)$$

por el Teorema Fundamental del Cálculo.

Luego, vale que

$$v_t(x, t) + \nabla_x v(x, t) \cdot \dot{\mathbf{x}}(t) + r(x, a) \leq 0.$$

Como  $\mathbf{x}(\cdot)$  es solución del sistema (2.2.1) para el control constante  $a$ , reemplazando  $\dot{\mathbf{x}}(\cdot)$  por  $f(\mathbf{x}(\cdot), a)$  se tiene

$$v_t(x, t) + f(x, a) \cdot \nabla_x v(x, t) + r(x, a) \leq 0,$$

y vale para todo  $a \in A$ , de donde concluimos que

$$v_t(x, t) + \max_{a \in A} \{ \nabla_x v(x, t) \cdot \dot{x}(t) + r(x, a) \} \leq 0. \quad (2.2.10)$$

Supongamos ahora que conocemos el control óptimo  $\alpha^*(\cdot)$  y la solución  $x^*(\cdot)$  del problema.

Por el Principio de Programación Dinámica, y utilizando  $\alpha^*(\cdot)$  para los tiempos  $t \leq s < t+h$ ,

$$\begin{aligned} v(x, t) &= \int_t^T r(\mathbf{x}^*(s), \alpha^*(s)) ds + g(\mathbf{x}^*(T)) \\ &= \int_t^{t+h} r(\mathbf{x}^*(s), \alpha^*(s)) ds + v(\mathbf{x}^*(t+h), t+h). \end{aligned}$$

Si reordenamos como antes y dividimos por  $h > 0$ ,

$$\frac{v(\mathbf{x}^*(t+h), t+h) - v(\mathbf{x}^*(t), t)}{h} + \frac{1}{h} \int_t^{t+h} r(\mathbf{x}^*(s), \alpha^*(s)) ds = 0,$$

y haciendo la misma cuenta que antes, con  $h \rightarrow 0$ ,

$$v_t(\mathbf{x}^*(t), \alpha^*(t)) + \nabla_x v(\mathbf{x}^*(t), t) \cdot \dot{\mathbf{x}}^*(t) + r(\mathbf{x}^*(t), \alpha^*(t)) = 0,$$

pero además,  $\mathbf{x}^*(t) = x$ ,  $\dot{\mathbf{x}}^*(t) = f(x, \alpha^*(t))$ .

Definiendo  $a^* = \alpha^*(t)$ , con  $a^* \in A$ ,

$$v_t(x, t) + \nabla_x v(x, t) \cdot f(x, a^*) + r(x, a^*) = 0$$

para  $a^* \in A$ . Esto es, vale la igualdad en (2.2.10), y queda demostrado el teorema.  $\square$

*Observación 2.2.1.* Las ecuaciones llamadas de Hamilton-Jacobi-Bellman son de la forma

$$v_t + H(x, \nabla_x v) = 0.$$

En nuestro caso,

$$H(x, \nabla_x v) := \max_{a \in A} H(x, \nabla_x v, a) = \max_{a \in A} \{f(x, a) \cdot \nabla_x v(x, t) + r(x, t)\}.$$

En general, utilizaremos la variable  $p$  para el gradiente de  $v$ , y escribiremos  $H(x, p)$  para referirnos al Hamiltoniano  $H$ .

### 2.2.3. Diseño del Control Óptimo

Utilizaremos también el Método de Programación Dinámica para diseñar los controles óptimos de la siguiente forma:

1. Resolvemos la ecuación de Hamilton-Jacobi-Bellman (2.2.9) para obtener  $v$ .
2. Una vez conocido  $v$ , para cada  $x \in \mathbb{R}^n$  y cada  $0 \leq t \leq T$  definimos  $\alpha(x, t) = a \in A$  con  $a$  el parámetro en donde se alcanza el máximo en (2.2.9) para  $(x, t)$ . Es decir, elegimos  $\alpha(x, t)$  tal que

$$v_t(x, t) + f(x, \alpha(x, t)) \cdot \nabla_x v(x, t) + r(x, \alpha(x, t)) = 0.$$

3. Resolvemos el siguiente sistema, suponiendo que  $\alpha(x, t)$  es suficientemente regular:

$$\begin{cases} \dot{\mathbf{x}}^*(s) &= f(\mathbf{x}^*(s), \alpha(\mathbf{x}^*(s), s)) & (t \leq s \leq T) \\ \mathbf{x}^*(t) &= x. \end{cases} \quad (2.2.11)$$

4. Definimos el control óptimo como

$$\alpha^*(s) = \alpha(\mathbf{x}^*(s), s). \quad (2.2.12)$$

En resumen, si el estado del sistema a tiempo  $t$  es  $x$ , usamos el control que a tiempo  $t$  toma el valor  $a \in A$  tal que en  $a$  se alcanza el máximo de Hamilton-Jacobi-Bellman.

Probemos ahora que el control construido es, efectivamente, óptimo.

**Teorema 2.2.3** (Verificación de la Optimalidad). *El control  $\alpha^*(\cdot)$  definido por (2.2.12) es óptimo.*

*Demostración.* Se tiene que

$$P_{x,t}[\alpha^*(\cdot)] = \int_t^T r(\mathbf{x}^*(s), \alpha^*(s)) ds + g(\mathbf{x}^*(T)).$$

Por la construcción de  $\alpha^*(\cdot)$ , podemos reemplazar  $r$  y obtenemos

$$\begin{aligned} P_{x,t}[\alpha^*(\cdot)] &= \int_t^T (-v_t(\mathbf{x}^*(s), s) - f(\mathbf{x}^*(s), \alpha^*(s)) \cdot \nabla_x v(\mathbf{x}^*(s), s) + g(\mathbf{x}^*(T))) \\ &= - \int_t^T (v_t(\mathbf{x}^*(s), s) + \nabla_x v(\mathbf{x}^*(s), s) \cdot \dot{\mathbf{x}}^*(s)) ds + g(\mathbf{x}^*(T)) \\ &= - \int_t^T \frac{d}{ds} v(\mathbf{x}^*(s), s) ds + g(\mathbf{x}^*(T)) \\ &= -v(\mathbf{x}^*(T), T) + v(\mathbf{x}^*(t), t) + g(\mathbf{x}^*(T)) = v(x, t). \end{aligned}$$

Luego, como

$$v(x, t) = \sup_{\alpha(\cdot) \in \mathcal{A}} P_{x,t}[\alpha(\cdot)].$$

se tiene

$$P_{x,t}[\alpha^*(\cdot)] = \sup_{\alpha(\cdot) \in \mathcal{A}} P_{x,t}[\alpha(\cdot)];$$

lo que demuestra que  $\alpha^*(\cdot)$  es el óptimo.  $\square$

*Observación 2.2.2.* En la demostración anterior, usamos la unicidad de la solución de Hamilton-Jacobi-Bellman, afirmación que todavía no probamos.

## 2.3. Soluciones Viscosas

En la sección anterior mostramos que podíamos encontrar un control óptimo valiéndonos de la solución de cierta ecuación diferencial. Dicha solución no tiene por qué existir, al menos en el sentido clásico. En esta sección definiremos un concepto de

soluciones débiles para ecuaciones del tipo Hamilton-Jacobi. Estas son las soluciones viscosas. Mostraremos que son consistentes con las soluciones clásicas y que, de existir, son únicas. Luego, se deduce una fórmula que caracteriza estas soluciones cuando el Hamiltoniano depende sólo del gradiente, y por último, veremos que la función valor es la (única) solución viscosa de la ecuación de Hamilton-Jacobi-Bellman de control óptimo.

**Definición 2.3.1.** *Una ecuación de Hamilton-Jacobi tiene la forma*

$$\begin{cases} u_t + H(\nabla_x u, \mathbf{x}) = 0 & \text{en } \mathbb{R}^n \times (0, \infty) \\ u = g & \text{en } \mathbb{R}^n \times \{t = 0\} \end{cases} \quad (2.3.1)$$

Aquí, el Hamiltoniano  $H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ , y el dato inicial  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  son conocidos. La incógnita es  $u = u(x, t)$ ,  $u : \mathbb{R}^n \times [0, \infty) \rightarrow \mathbb{R}$ .

En algunos casos escribiremos  $H = H(p, x)$ , e indicamos con  $p$  la variable que corresponde a  $\nabla_x u$ .

**Definición 2.3.2.** *Una función  $u$  acotada y uniformemente continua se dice solución viscosa del problema a valores iniciales (2.3.1) si*

i)  $u = g$  en  $\mathbb{R}^n \times \{t = 0\}$ ,

ii) para cada  $v \in C^\infty(\mathbb{R}^n \times (0, \infty))$  vale

a) Si  $u - v$  tiene un máximo local en  $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$ , entonces

$$v_t(x_0, t_0) + H(\nabla_x v(x_0, t_0), x_0) \leq 0$$

b) Si  $u - v$  tiene un mínimo local en  $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$ , entonces

$$v_t(x_0, t_0) + H(\nabla_x v(x_0, t_0), x_0) \geq 0$$

Si sólo verifica a), la llamamos subsolución, y si sólo verifica b), la llamamos supersolución.

### 2.3.1. Consistencia

Veamos que la definición de solución viscosa es consistente con la de solución clásica.

**Lema 2.3.1.** *Si  $u \in C^1(\mathbb{R}^n \times (0, \infty))$  es solución de (2.3.1), acotada y uniformemente continua, entonces es solución viscosa.*

*Demostración.* Si  $v$  es suave, y  $u - v$  tiene un máximo local en  $(x_0, t_0)$ ,

$$\nabla_x u(x_0, t_0) = \nabla_x v(x_0, t_0)$$

$$u_t(x_0, t_0) = v_t(x_0, t_0),$$

entonces

$$u_t(x_0, t_0) + H(\nabla_x u(x_0, t_0), x_0) = v_t(x_0, t_0) + H(\nabla_x u(x_0, t_0), x_0) = 0$$

dado que  $u$  es solución de (2.3.1).

De la misma forma, podemos probar que vale la igualdad para cualquier  $v$  tal que  $u - v$  tenga un mínimo local en  $(x_0, t_0)$ . □

Ahora queremos ver que una solución viscosa suficientemente regular, es solución clásica. Para ello, usaremos el siguiente lema:

**Lema 2.3.2.** *Sea  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  continua y diferenciable en  $y_0$ . Entonces existe  $v \in C^1(\mathbb{R}^n)$  tal que  $u(y_0) = v(y_0)$  y  $u - v$  tiene máximo local estricto en  $y_0$ .*

La demostración de este lema se puede ver en [2], pag. 544-545.

**Teorema 2.3.1.** *Sea  $u$  una solución viscosa de (2.3.1),  $u$  diferenciable en  $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$ . Entonces,*

$$u_t(x_0, t_0) + H(\nabla_x u(x_0, t_0), x_0) = 0.$$

*Demostración.* Dividamos la demostración en distintos pasos.

1. Primero usamos el Lema anterior para  $u$ , con  $\mathbb{R}^{n+1}$  reemplazando a  $\mathbb{R}^n$ , y  $(x_0, t_0) = y_0$ . Entonces existe  $v \in C^1(\mathbb{R}^{n+1})$  tal que  $u - v$  tiene un máximo local estricto en  $(x_0, t_0)$ .
2. Sea  $v^\varepsilon := \eta^\varepsilon * v$ , con  $\eta^\varepsilon \in C_0^\infty(B(0, \varepsilon))$  una función regularizante, entonces se tiene

$$v^\varepsilon \rightarrow v,$$

$$\nabla_x v^\varepsilon \rightarrow \nabla_x v,$$

$$v_t^\varepsilon \rightarrow v_t,$$

uniformemente en un entorno de  $(x_0, t_0)$ .

Como  $u - v$  tiene un máximo local estricto,  $u - v^\varepsilon$  tiene un máximo en algún  $(x_\varepsilon, t_\varepsilon)$ , y deducimos de la convergencia uniforme que, cuando  $\varepsilon \rightarrow 0$ , se tiene que  $(x_\varepsilon, t_\varepsilon) \rightarrow (x_0, t_0)$ .

Por la definición de solución viscosa,

$$v_t^\varepsilon(x_\varepsilon, t_\varepsilon) + H(\nabla_x v^\varepsilon(x_\varepsilon, t_\varepsilon), x_\varepsilon) \leq 0.$$

Tomando límite con  $\varepsilon \rightarrow 0$ ,

$$v_t(x_0, t_0) + H(\nabla_x v(x_0, t_0), x_0) \leq 0.$$

Ahora, por lo que hicimos en la parte (1) de la demostración, y como  $u$  es diferenciable en  $(x_0, t_0)$  tiene que valer

$$\nabla_x u(x_0, t_0) = \nabla_x v(x_0, t_0),$$

$$u_t(x_0, t_0) = v_t(x_0, t_0),$$

de donde deducimos que

$$u_t(x_0, t_0) + H(\nabla_x u(x_0, t_0), x_0) \leq 0.$$

3. Si aplicamos ahora el lema previo a  $-u$ , en  $\mathbb{R}^{n+1}$ , obtenemos  $v \in C^1(\mathbb{R}^{n+1})$  tal que  $u - v$  tiene un mínimo local estricto. Emulando el argumento anterior, se tiene

$$u_t(x_0, t_0) + H(\nabla_x u(x_0, t_0), x_0) \geq 0.$$

lo que concluye la demostración.

□

### 2.3.2. Unicidad

Fijemos ahora  $T > 0$  y estudiemos el problema

$$\begin{cases} u_t + H(\nabla_x u, t) = 0 & \text{en } \mathbb{R}^n \times (0, T] \\ u = g & \text{en } \mathbb{R}^n \times \{t = 0\} \end{cases} \quad (2.3.2)$$

Diremos que una función  $u$  acotada y uniformemente continua es una solución viscosa de (2.3.2) si

i)  $u = g$  en  $\mathbb{R}^n \times \{t = 0\}$ ,

ii) para cada  $v \in C^\infty(\mathbb{R}^n \times (0, \infty))$  vale:

a) si  $u - v$  tiene un máximo local en  $(x_0, t_0) \in \mathbb{R}^n \times (0, T]$ , entonces

$$v_t(x_0, t_0) + H(\nabla_x v(x_0, t_0), x_0) \leq 0.$$

b) si  $u - v$  tiene un mínimo local en  $(x_0, t_0) \in \mathbb{R}^n \times (0, T]$ , entonces

$$v_t(x_0, t_0) + H(\nabla_x v(x_0, t_0), x_0) \geq 0.$$

Es decir, también permitimos  $t_0 = T$ .

Se tiene el siguiente resultado:

**Teorema 2.3.2.** (*Unicidad*)

Supongamos que existe una constante positiva  $c$  tal que para todo  $x, y, p, q \in \mathbb{R}^n$ ,  $H$  satisface las siguientes condiciones:

$$|H(p, x) - H(q, x)| \leq c|p - q| \quad (2.3.3)$$

$$|H(p, x) - H(p, y)| \leq c|x - y|(1 + |p|) \quad (2.3.4)$$

entonces existe a lo sumo una solución viscosa de (2.3.2).

Vamos a omitir la demostración de este teorema por ser muy extensa, pero utiliza una técnica muy útil consistente en duplicar el número de variables, ver [2] pag. 547-550.

### 2.3.3. Fórmula de Lax-Hopf

En las secciones anteriores estudiamos el problema de maximizar un pago. En esta sección, mostraremos resultados para el problema de minimizar un costo, que es equivalente al anterior. Cuando el Hamiltoniano  $H$  es convexo y depende sólo de  $p$ , tenemos una fórmula explícita para la solución del problema

$$\begin{cases} u_t + H(\nabla_x u) = 0 & \text{en } \mathbb{R}^n \times (0, T] \\ u = g & \text{en } \mathbb{R}^n \times \{t = 0\}. \end{cases} \quad (2.3.5)$$

Supongamos que  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  es Lipchitz,  $p \mapsto H(p)$  es convexa, y que

$$\lim_{|p| \rightarrow \infty} \frac{H(p)}{|p|} = +\infty.$$

Entonces, la fórmula de Lax-Hopf para una solución de (2.3.5) es

$$u(x, t) = \min_{y \in \mathbb{R}^n} \left\{ tL \left( \frac{x - y}{t} \right) + g(y) \right\} \quad (2.3.6)$$

para  $x \in \mathbb{R}^n$ ,  $t \in [0, T]$ , siendo  $L$  la transformada de Legendre de  $H$

$$L(q) = \sup_{p \in \mathbb{R}^n} \{p \cdot q - H(p)\}, \quad (q \in \mathbb{R}^n)$$

**Teorema 2.3.3.** *Bajo las hipótesis anteriores, si además  $g$  es acotada, entonces la única solución viscosa de (2.3.5) está dada por (2.3.6).*

Vamos a omitir la demostración de este teorema por ser un caso particular del teorema que veremos a continuación, y su demostración puede adaptarse sin mucha dificultad, ver [2] Capítulos 3 y 10.

Demostremos que la función valor definida para el problema de control es la única solución viscosa de la ecuación HJB deducida antes.

**Teorema 2.3.4.** *La función valor  $v$  es la única solución viscosa del problema a valores terminales*

$$\begin{cases} v_t + \min_{a \in A} \{ \mathbf{f}(x, a) \cdot \nabla_x v(x, t) + r(x, a) \} = 0 & \mathbb{R}^n \times (0, T) \\ v = g & \mathbb{R}^n \times \{t = T\} \end{cases} \quad (2.3.7)$$

*Observación 2.3.1.* Como tenemos un problema a valores terminales, debemos especificar a qué llamamos solución viscosa en este contexto: una función  $v$  acotada, uniformemente continua es solución viscosa de (2.3.7) si

- i)  $v = g$  en  $\mathbb{R}^n \times \{t = T\}$ ,
- ii) para cada  $u \in C^\infty(\mathbb{R}^n \times (0, T))$  vale
  - a) Si  $v - u$  tiene un máximo local en  $(x_0, t_0) \in \mathbb{R}^n \times (0, T)$ , entonces
 
$$u_t(x_0, t_0) + H(\nabla_x u(x_0, t_0), x_0) \geq 0.$$
  - b) Si  $v - u$  tiene un mínimo local en  $(x_0, t_0) \in \mathbb{R}^n \times (0, T)$ , entonces
 
$$u_t(x_0, t_0) + H(\nabla_x u(x_0, t_0), x_0) \leq 0.$$

Es decir, cambiamos el sentido de las desigualdades.

*Observación 2.3.2.* Si  $v$  es solución viscosa de la ecuación (2.3.7) entonces  $w(x, t) = v(x, T - t)$  es solución viscosa de

$$\begin{cases} w_t - \min_{a \in A} \{ \mathbf{f}(x, a) \cdot \nabla_x w(x, t) + r(x, a) \} = 0 & \mathbb{R}^n \times (0, T) \\ w = g & \mathbb{R}^n \times \{t = 0\} \end{cases}$$

*Demostración del Teorema 2.3.4.* Supongamos que  $v - u$  tiene un mínimo local en  $(x_0, t_0) \in \mathbb{R}^n \times (0, T)$ . Queremos mostrar que

$$u_t(x_0, t_0) + \min_{a \in A} \{ \mathbf{f}(x_0, a) \cdot \nabla_x u(x_0, t_0) + r(x_0, a) \} \leq 0$$

Supongamos que no es cierto, entonces existe  $\theta > 0$  tal que

$$u_t(x, t) + \mathbf{f}(x, a) \cdot \nabla_x u(x, t) + r(x, a) \geq \theta > 0$$

para todo  $a \in A$  y  $(x, t)$  suficientemente cerca de  $(x_0, t_0)$ , o sea, para algún  $\delta > 0$

$$|x - x_0| + |t - t_0| < \delta \quad (2.3.8)$$

Como  $v - u$  tiene un mínimo local en  $(x_0, t_0)$ , podemos suponer

$$(v - u)(x, t) \geq (v - u)(x_0, t_0)$$

para todo  $(x, t)$  que verifique (2.3.8).

Sea  $0 < h < \delta$  suficientemente chico tal que  $|\mathbf{x}(s) - x_0| < \delta$  para todo  $t_0 \leq s \leq t_0 + h$ , donde  $\mathbf{x}(\cdot)$  es la solución de

$$\begin{cases} \dot{\mathbf{x}}(s) = \mathbf{f}(\mathbf{x}(s), \alpha(s)) & t_0 < s < T \\ \mathbf{x}(t_0) = x_0 \end{cases}$$

para cierto control  $\alpha(\cdot) \in \mathbf{A}$ .

Entonces, para cualquier  $\alpha(\cdot) \in \mathbf{A}$  tenemos

$$\begin{aligned} v(\mathbf{x}(t_0 + h), t_0 + h) - v(x_0, t_0) &\geq u(\mathbf{x}(t_0 + h), t_0 + h) - u(x_0, t_0) \\ &= \int_{t_0}^{t_0+h} \frac{d}{ds} u(\mathbf{x}(s), s) ds \\ &= \int_{t_0}^{t_0+h} u_t(\mathbf{x}(s), s) \\ &\quad + \mathbf{f}(\mathbf{x}(s), \alpha(s)) \cdot \nabla_x u(\mathbf{x}(s), s) ds. \end{aligned}$$

Por otro lado, por la condición de optimalidad (recordemos que en este caso estamos mirando la ecuación con un mínimo!), podemos elegir  $\alpha(\cdot) \in \mathbf{A}$  tal que

$$v(x_0, t_0) \geq \int_{t_0}^{t_0+h} r(\mathbf{x}(s), \alpha(s)) ds + v(\mathbf{x}(t_0 + h), t_0 + h) - \frac{\theta h}{2}.$$

De las últimas dos desigualdades obtenemos

$$\frac{\theta h}{2} \geq \int_{t_0}^{t_0+h} u_t(\mathbf{x}(s), s) + \mathbf{f}(\mathbf{x}(s), \alpha(s)) \cdot \nabla_x u(\mathbf{x}(s), s) + r(\mathbf{x}(s), \alpha(s)) ds \geq \theta h,$$

lo cual es un absurdo.

De la misma manera se llega a un absurdo suponiendo que  $v - u$  tiene un máximo local en  $(x_0, t_0) \in \mathbb{R}^n \times (0, T)$  pero no satisface la desigualdad correspondiente.  $\square$

# Capítulo 3

## Control óptimo estocástico

### 3.1. El problema

En el capítulo anterior estudiamos el problema de control asociado a una ecuación diferencial ordinaria. En muchos casos, un mejor modelo para ciertos fenómenos es una ecuación diferencial estocástica:

$$\begin{cases} \dot{X}(t) = f(X(t)) + \sigma\xi(t), & t > 0 \\ X(0) = x_0 \end{cases} \quad (3.1.1)$$

donde  $\xi(\cdot)$  representa un término de *ruido blanco* que causa fluctuaciones aleatorias. Observemos que en este caso la solución también será aleatoria. Usaremos  $X(\cdot)$  mayúscula para indicar esta característica.

Una solución de (3.1.1) es un proceso estocástico, con una distribución de probabilidad sobre los posibles caminos.

Podríamos entonces, considerar el problema de control asociado a una ecuación diferencial estocástica.

Sea ahora  $f : \mathbb{R}^n \times A \rightarrow \mathbb{R}^n$  y el siguiente sistema

$$\begin{cases} \dot{\mathbf{X}}(s) = f(\mathbf{X}(s), \mathbf{A}(s)) + \sigma\xi(s) & t \leq s \leq T \\ \mathbf{X}(t) = x_0. \end{cases} \quad (3.1.2)$$

Introduzcamos las siguientes definiciones:

**Definición 3.1.1.** *Un control  $\mathbf{A}(\cdot)$  es una asignación de  $[t, T]$  en  $A$ , tal que para cada tiempo  $t \leq s \leq T$ ,  $\mathbf{A}(s)$  depende sólo de  $s$  y de las observaciones de  $\mathbf{X}(\tau)$  para  $t \leq \tau \leq s$ .*

**Definición 3.1.2.** *El funcional de pago asociado es*

$$P_{x,t}[\mathbf{A}(\cdot)] = \mathbb{E} \left\{ \int_t^T r(\mathbf{X}(s), \mathbf{A}(s)) ds + g(\mathbf{X}(T)) \right\}, \quad (3.1.3)$$

*esto es, el valor esperado sobre todos los caminos de la solución de (3.1.2). Llamamos  $r$  al pago variable y  $g$  al pago terminal, y las consideramos funciones conocidas.*

El problema básico será encontrar un control óptimo  $A^*(\cdot)$ , tal que

$$P_{x,t}[A^*(\cdot)] = \max_{A(\cdot)} P_{x,t}[A(\cdot)].$$

En la Sección §2 haremos una breve introducción a las ecuaciones y al cálculo estocástico. En la Sección §3, adaptaremos el método de programación dinámica a este modelo, y calcularemos, de manera heurística, la ecuación de Hamilton-Jacobi-Bellman correspondiente; por último, diseñaremos un control óptimo. Se puede verificar que la función valor es la solución viscosa de la ecuación que encontramos, para más detalles, ver [3].

## 3.2. Ecuaciones y cálculo estocástico

Introduzcamos algunas definiciones necesarias. Como antes, continuaremos usando negritas para denotar funciones vectoriales y también procesos estocásticos vectoriales.

**Definición 3.2.1.** *Un proceso estocástico es una colección de variables aleatorias  $X(t)$ ,  $0 \leq t < \infty$ , cada una definida en el mismo espacio de probabilidad  $(\Omega, \mathcal{F}, P)$ .*

La asignación  $t \mapsto X(t, \omega)$  es el  $\omega$ -ésimo camino del proceso.

**Definición 3.2.2.** *Un proceso estocástico  $W(t)$  que toma valores en  $\mathbb{R}$  se llama un movimiento browniano unidimensional si*

- (i)  $W(0) = 0$  casi seguramente (es decir, excepto un conjunto de probabilidad nula),
- (ii) casi todo camino es continuo. Esto es, existe un conjunto  $\Omega_0 \subset \Omega$ , con  $P(\Omega_0) = 1$ , tal que para todo  $\omega \in \Omega_0$ , tenemos que  $t \mapsto W(t, \omega)$  es continua.
- (iii)  $W(t)$  tiene distribución normal con  $\mu = 0$ ,  $\sigma^2 = t$ , es decir,

$$W(t) \sim \mathcal{N}(0, t),$$

- (iv) para toda elección de tiempos  $0 < t_1 < t_2 < \dots < t_m$ , las variables aleatorias

$$W(t_1), W(t_2) - W(t_1), \dots, W(t_m) - W(t_{m-1})$$

son independientes.

Debido a la propiedad (iv), decimos que  $W$  tiene *incrementos independientes*.

Podemos interpretar el ruido blanco unidimensional  $\xi(\cdot)$  como la derivada temporal de un movimiento browniano,

$$\xi = \frac{dW(t)}{dt}.$$

Sin embargo, esta es sólo una notación conveniente, y no debemos olvidar que el camino  $t \mapsto X(t, \omega)$  no es diferenciable en ningún punto para casi todo  $\omega$ .

**Definición 3.2.3.** *Un movimiento browniano  $n$ -dimensional es*

$$\mathbf{W}(t) = (W^1(t), W^2(t), \dots, W^n(t))^T$$

donde los  $W^i(t)$  son movimientos brownianos unidimensionales independientes.

Volvamos a considerar el sistema de ecuaciones estocásticas (sin el control  $A$ ):

$$\begin{cases} \dot{\mathbf{X}}(t) = f(\mathbf{X}(t)) + \sigma \xi(t) & (t > 0) \\ \mathbf{X}(0) = x_0 \end{cases} \quad (3.2.1)$$

y aceptemos, informalmente, que  $\xi = \dot{\mathbf{W}}$ .

**Definición 3.2.4.** *Decimos que un proceso estocástico  $\mathbf{X}(\cdot)$  es una solución de la ecuación (3.2.1) si para todo tiempo  $t \geq 0$ , se tiene*

$$\mathbf{X}(t) = x_0 + \int_0^t f(\mathbf{X}(s)) ds + \sigma \mathbf{W}(t). \quad (3.2.2)$$

*Observación 3.2.1.* Es posible resolver localmente la ecuación (3.2.2) por el método de aproximaciones sucesivas, asumiendo cierta regularidad de  $f$ , por ejemplo, que sea continua Lipchitz.

Partiendo de  $\mathbf{X}^0(\cdot) = x_0$ , definimos inductivamente

$$\mathbf{X}^{k+1}(\cdot) := x_0 + \int_0^t \mathbf{f}(\mathbf{X}^k(s)) ds + \sigma \mathbf{W}(t).$$

Podemos demostrar que es una sucesión de Cauchy de manera muy similar al caso de ecuaciones diferenciales ordinarias, y tenemos que  $\mathbf{X}^k(t)$  converge a  $\mathbf{X}(t)$  para todo  $t \geq 0$ , y que  $\mathbf{X}(\cdot)$  es una solución de (3.2.2), para más detalles ver [4] pag. 83-84.

*Observación 3.2.2.* En el capítulo 5 consideraremos ecuaciones estocásticas más generales:

$$\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{X}(t)) + \mathbf{H}(\mathbf{X}(t))\xi(t) \quad (t > 0), \quad (3.2.3)$$

que podemos escribir, sólo por notación, como antes,

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{f}(\mathbf{X}(t)) + \mathbf{H}(\mathbf{X}(t)) \frac{d\mathbf{W}(t)}{dt}.$$

Reformulando esta ecuación como

$$d\mathbf{X}(t) = \mathbf{f}(\mathbf{X}(t))dt + \mathbf{H}(\mathbf{X}(t))d\mathbf{W}(t),$$

nos queda una *ecuación diferencial estocástica de Itô*. Diremos que  $\mathbf{X}(\cdot)$  es su solución, dada la condición inicial  $\mathbf{X}(0) = x_0$ , si

$$\mathbf{X}(t) = x_0 + \int_0^t f(\mathbf{X}(s)) ds + \int_0^t \mathbf{H}(\mathbf{X}(s)) \cdot d\mathbf{W}(s)$$

para todo tiempo  $t \geq 0$ . En esta expresión,  $\int_0^t \mathbf{H}(\mathbf{X}) \cdot d\mathbf{W}$  se dice una *integral estocástica de Ito*.

En general, dados un movimiento Browniano  $\mathbf{W}(\cdot)$  y un proceso  $\mathbf{Y}(\cdot)$  tal que para cada tiempo  $0 \leq s \leq t$ ,  $\mathbf{Y}(s)$  depende sólo de  $\mathbf{W}(\tau)$  para tiempos  $0 \leq \tau \leq s$  (y que no depende de  $\mathbf{W}(\tau)$  para  $\tau > s$ ), se puede definir la integral estocástica de Itô

$$\int_0^t \mathbf{Y} \cdot d\mathbf{W}.$$

Si bien no entraremos en detalles sobre la construcción de la integral de Itô, una de sus propiedades más importantes es la siguiente:

$$E \left[ \int_0^t \mathbf{Y} \cdot d\mathbf{W} \right] = 0. \quad (3.2.4)$$

### 3.2.1. Regla de la Cadena de Itô

Una vez definida la integral estocástica de Itô, se tiene un nuevo cálculo, cuyas propiedades deberíamos estudiar. En particular, la regla de la cadena ahora contiene términos adicionales en comparación a la regla de la cadena usual. Estos cambios influirán en la deducción de la ecuación de Hamilton-Jacobi-Bellman.

Enunciemos el siguiente resultado, cuya demostración puede verse en [4].

**Teorema 3.2.1** (Caso unidimensional). *Sea  $u : \mathbb{R} \rightarrow \mathbb{R}$  de clase  $C^3$ . Dada la ecuación*

$$\begin{cases} dX(t) = F(t)dt + G(t)dW(t) & t \geq 0 \\ X(0) = x_0. \end{cases} \quad (3.2.5)$$

y su solución

$$X(t) = x_0 + \int_0^t F(t)dt + \int_0^t G(t)dW(t),$$

definimos el proceso

$$Y(t) := u(X(t)).$$

Entonces,

$$dY = \left[ u'(X(t))F(t) + \frac{1}{2}G^2(t)u''(X(t)) \right] dt + u'(X(t))G(t)dW(t).$$

Para dar una idea de la demostración del teorema, debemos utilizar el siguiente principio heurístico:  $dW(t) = (dt)^{1/2}$ .

Ahora, veamos qué ecuación resuelve  $Y(t)$  en el tiempo.

Aproximemos primero la función  $u$  con su polinomio de Taylor de segundo grado, incluyendo una estimación del error,

$$u(s + ds) = u(s) + u'(s)ds + \frac{1}{2}u''(s)(ds)^2 + o([ds]^3).$$

Entonces,

$$\begin{aligned} dY(t) &= d(u(X(t))) \\ &= u'(X(t))dX(t) + \frac{1}{2}u''(X(t))dX(t)^2 + o(dX(t)^3) \\ &= u'(X(t))[F(t)dt + G(t)dW(t)] \\ &\quad + \frac{1}{2}u''(X(t))[F(t)dt + G(t)dW(t)]^2 + o([dt + dW]^2) \end{aligned}$$

donde hemos reemplazado  $dX$  en la última igualdad utilizando la ecuación (3.2.5).

Ahora, usando  $dW(t) = (dt)^{1/2}$ ,

$$\begin{aligned} [F(t)dt + G(t)dW(t)]^2 &= F(t)^2dt^2 + 2F(t)G(t)dt dW(t) + G(t)^2dW(t)^2 \\ &= G(t)^2dt + o(dt). \end{aligned}$$

Entonces, ignorando los términos de orden  $o(dt)$ , llegamos a la regla de la cadena de Itô unidimensional:

$$\begin{aligned} dY(t) &= d(u(X(t))) \\ &= \left[ u'(X(t))F(t) + \frac{1}{2}G^2(t)u''(X(t)) \right] dt + u'(X(t))G(t)dW(t). \end{aligned}$$

*Observación 3.2.3.* Reescribiendo  $u$  en forma integral, tenemos para todo  $t \geq 0$ ,

$$\begin{aligned} u(X(t)) &= Y(t) \\ &= Y(0) + \int_0^t \left[ u'(X(s))F(s) + \frac{1}{2}G^2(s)u''(X(s)) \right] ds \\ &\quad + \int_0^t u'(X(s))G(s)dW(s), \end{aligned}$$

donde  $X(t)$  era la solución de la ecuación (3.2.5).

*Observación 3.2.4.* De manera similar, tenemos una versión para dimensiones superiores. Por simplicidad consideraremos solamente el siguiente caso:

$$\begin{cases} d\mathbf{X}(t) &= \mathbf{F}(t)dt + \sigma d\mathbf{W}(t) \quad (t \geq 0) \\ \mathbf{X}(0) &= x_0. \end{cases} \quad (3.2.6)$$

Aquí,  $\mathbf{X}(t) = (X^1(t), \dots, X^n(t))^T$ .

Este sistema estocástico dice que para cada índice  $i$ , se tiene

$$dX^i(t) = F^i(t)dt + \sigma dW^i(t).$$

Sea  $u : \mathbb{R}^n \times [0, \infty) \rightarrow \mathbb{R}$  de clase  $C^3$ , y definamos

$$Y(t) := u(\mathbf{X}(t), t).$$

Como antes, calculamos  $dY(t)$

$$\begin{aligned} dY(t) &= d[u(\mathbf{X}(t))] \\ &= u_t(\mathbf{X}(t), t)dt + \sum_{i=1}^n u_{x_i}(\mathbf{X}(t), t)dX^i(t) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n u_{x_i x_j}(\mathbf{X}(t), t)dX^i(t)dX^j(t). \end{aligned}$$

Necesitamos utilizar que  $\mathbf{X}(\cdot)$  es solución de (3.2.6), que  $dW^i = (dt)^{\frac{1}{2}}$ , y que

$$dW^i dW^j = \begin{cases} dt & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

esto último porque las componentes de  $d\mathbf{W}$  son independientes.

Reemplazando estas identidades en la cuenta anterior, y quedándonos sólo con los términos de primer y segundo orden en  $dt$ , tenemos:

$$\begin{aligned} dY(t) &= u_t(\mathbf{X}(t), t)dt + \sum_{i=1}^n u_{x_i}(\mathbf{X}(t), t)[F^i(t)dt + \sigma dW^i(t)] \\ &\quad + \frac{\sigma^2}{2} \sum_{i=1}^n u_{x_i x_i}(\mathbf{X}(t), t)dt. \end{aligned}$$

Observemos que la primer sumatoria involucra el producto interno de  $\nabla_x u$  con  $\mathbf{F}(t)dt + \sigma d\mathbf{W}(t)$ , mientras que la segunda es el laplaciano de  $u$ . Podemos escribir en forma más compacta

$$dY(t) = u_t(\mathbf{X}, t) + \nabla_x u(\mathbf{X}, t) \cdot [\mathbf{F}(t)dt + \sigma d\mathbf{W}(t)] + \frac{\sigma^2}{2} \Delta u(\mathbf{X}, t)dt.$$

Esta es la regla de la cadena de Itô  $n$ -dimensional, que también involucra derivadas segundas como en el caso unidimensional.

### 3.3. Programación Dinámica

Queremos, emulando lo hecho para control determinístico, diseñar un control óptimo para el caso estocástico. Necesitamos ver qué significa el principio de Programación Dinámica en este contexto.

Para ello, volvamos a considerar el problema de control para el siguiente sistema de ecuaciones estocásticas

$$\begin{cases} d\mathbf{X}(s) = \mathbf{f}(\mathbf{X}(s), \mathbf{A}(s))ds + \sigma d\mathbf{W}(s) & t \leq s \leq T \\ \mathbf{X}(t) = x. \end{cases} \quad (3.3.1)$$

Por lo estudiado antes, una solución tendrá la forma

$$\mathbf{X}(\tau) = x + \int_0^\tau \mathbf{f}(\mathbf{X}(s), \mathbf{A}(s))ds + \sigma[\mathbf{W}(\tau) - \mathbf{W}(t)]$$

para todo  $t \leq \tau \leq T$ .

Recordemos la definición del *funcional de pago esperado*.

$$P_{x,t}[\mathbf{A}(\cdot)] := \mathbb{E} \left\{ \int_t^\tau r(\mathbf{X}(s), \mathbf{A}(s))ds + g(\mathbf{X}(t)) \right\}, \quad (3.3.2)$$

y la *función valor* será

$$v(x, t) := \sup_{\mathbf{A}(\cdot) \in \mathcal{A}} P_{x,t}[\mathbf{A}(\cdot)].$$

Vamos a aplicar el método de programación dinámica. Como hicimos antes, queremos encontrar una ecuación en derivadas parciales que verifique  $v$ , y luego usar esta ecuación para diseñar el control óptimo  $\mathbf{A}^*(\cdot)$ .

#### 3.3.1. Ecuación de Hamilton-Jacobi-Bellman

Como en el caso determinístico, queremos transformar el problema de control en otro problema, resolver una ecuación diferencial de la cual la función valor  $v$  sea solución.

Vamos a suponer que existe un control óptimo  $\mathbf{A}^*(\cdot)$ , si bien podríamos proceder como en el capítulo anterior, y demostrar su existencia trabajando con controles que están a menos de  $\varepsilon$  de realizar el supremo.

Sea entonces  $\mathbf{A}(\cdot)$  un control cualquiera, y supongamos que lo usamos para tiempos  $t \leq s \leq t+h$ ,  $h > 0$ , y a partir de ahí, usamos el control óptimo  $\mathbf{A}^*(\cdot)$ . Definimos, para  $\mathbf{A}(\cdot)$  y  $h$  fijos, el control admisible

$$\bar{\mathbf{A}}(s) := \begin{cases} \mathbf{A}(s) & \text{si } t \leq s \leq t+h \\ \mathbf{A}^*(s) & \text{si } t+h < s \leq T \end{cases}$$

Calculemos el pago esperado

$$\mathbb{E} \left[ \int_t^T r(\mathbf{X}(s), \bar{\mathbf{A}}(s)) ds + g(\mathbf{X}(t)) \right]$$

usando este control. Por la linealidad de la integral y de la esperanza, tenemos

$$\begin{aligned} & \mathbb{E} \left[ \int_t^{t+h} r(\mathbf{X}(s), \mathbf{A}(s)) ds + \int_{t+h}^T r(\mathbf{X}(s), \mathbf{A}^*(s)) ds + g(\mathbf{X}(t)) \right] = \\ & \mathbb{E} \left[ \int_t^{t+h} r(\mathbf{X}(s), \mathbf{A}(s)) ds \right] + \mathbb{E} \left[ \int_{t+h}^T r(\mathbf{X}(s), \mathbf{A}^*(s)) ds + g(\mathbf{X}(t)) \right] = \\ & \mathbb{E} \left[ \int_t^{t+h} r(\mathbf{X}(s), \mathbf{A}(s)) ds \right] + v(\mathbf{X}(t+h), t+h), \end{aligned}$$

y en la última igualdad utilizamos que  $\mathbf{A}^*(\cdot)$  es el control óptimo.

Pero como  $v(x, t)$  es el supremo de los pagos sobre todos los controles admisibles, tenemos

$$\mathbb{E} \left[ \int_t^{t+h} r(\mathbf{X}(s), \mathbf{A}(s)) ds \right] + v(\mathbf{X}(t+h), t+h) \leq v(x, t). \quad (3.3.3)$$

*Observación 3.3.1.* Vale la igualdad cuando elegimos  $\mathbf{A}(\cdot) = \mathbf{A}^*(\cdot)$ , un control óptimo.

Por la desigualdad (3.3.3) tenemos, para un control arbitrario  $\mathbf{A}(\cdot)$ ,

$$\mathbb{E} \left[ \int_t^{t+h} r(\mathbf{X}(s), \mathbf{A}(s)) ds + v(\mathbf{X}(t+h), t+h) - v(x, t) \right] \leq 0,$$

es decir,

$$\mathbb{E} \left[ \int_t^{t+h} r ds \right] + \mathbb{E} [v(\mathbf{X}(t+h), t+h) - v(x, t)] \leq 0. \quad (3.3.4)$$

Recordemos la fórmula de Itô, y que  $\mathbf{X}(\cdot)$  es solución de la ecuación (3.3.1), con lo cual

$$\begin{aligned} dv(\mathbf{X}(s), s) &= v_t(\mathbf{X}(s), s) ds + \sum_{i=1}^n v_{x_i}(\mathbf{X}(s), s) dX^i(s) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n v_{x_i x_j}(\mathbf{X}(s), s) dX^i(s) dX^j(s) \\ &= v_t ds + \nabla_x v \cdot (\mathbf{f} ds + \sigma d\mathbf{W}(s)) + \frac{\sigma^2}{2} \Delta v ds. \end{aligned}$$

Esto nos dice que

$$\begin{aligned} v(\mathbf{X}(t+h), t+h) - v(\mathbf{X}(t), t) &= \int_t^{t+h} \left( v_t + \nabla_x v \cdot \mathbf{f} + \frac{\sigma^2}{2} \Delta v \right) ds \\ &\quad + \int_t^{t+h} \sigma \nabla_x v \cdot d\mathbf{W}(s). \end{aligned}$$

Calculando la esperanza, tenemos

$$\mathbb{E}[v(\mathbf{X}(t+h), t+h) - v(x, t)] = \mathbb{E} \left[ \int_t^{t+h} \left( v_t + \nabla_x v \cdot \mathbf{f} + \frac{\sigma^2}{2} \Delta v \right) ds \right].$$

Si sumamos en ambos miembros  $\mathbb{E}[\int_t^{t+h} f(\mathbf{X}(s), \mathbf{A}(s)) ds]$ , por la desigualdad (3.3.4) nos queda

$$\mathbb{E} \left[ \int_t^{t+h} \left( r + v_t + \nabla_x v \cdot \mathbf{f} + \frac{\sigma^2}{2} \Delta v \right) ds \right] \leq 0.$$

Si dividimos por  $h > 0$ , y tomamos límite  $h \rightarrow 0^+$ , nos queda

$$r(\mathbf{X}(t), \mathbf{A}(t)) + v_t(\mathbf{X}(t), t) + \mathbf{f}(\mathbf{X}(t), \mathbf{A}(t)) \cdot \nabla_x v(\mathbf{X}(t), t) + \frac{\sigma^2}{2} \Delta v(\mathbf{X}(t), t) \leq 0,$$

y recordando que  $\mathbf{X}(t) = x$ , si fijamos  $\mathbf{A}(t) = a \in A$ , se transforma en

$$r(x, a) + v_t(x, t) + \mathbf{f}(x, a) \cdot \nabla_x v(x, t) + \frac{\sigma^2}{2} \Delta v(x, t) \leq 0.$$

Esto vale para todo  $x, t$ , y  $a$ . Además, como vale la igualdad para el control óptimo, tenemos

$$\max_{a \in A} \left\{ v_t + \mathbf{f} \cdot \nabla_x v + \frac{\sigma^2}{2} \Delta v + r \right\} = 0.$$

*Observación 3.3.2.* La deducción anterior nos dice que la función valor  $v$  de nuestro problema de control estocástico es solución de la siguiente ecuación en derivadas parciales:

$$\begin{cases} v_t(x, t) + \frac{\sigma^2}{2} \Delta v(x, t) + \max_{a \in A} \{ f(x, a) \cdot \nabla_x v(x, t) + r(x, t) \} = 0 \\ v(x, T) = g(x). \end{cases} \quad (3.3.5)$$

La llamamos *Ecuación de Hamilton-Jacobi-Bellman* para el problema de control estocástico.

### 3.3.2. Diseño del Control Óptimo

Queremos ahora hallar el control óptimo  $\mathbf{A}^*(s)$ . Supongamos que podemos resolver la ecuación de Hamilton-Jacobi-Bellman para el caso estocástico, y en consecuencia, conocemos la función  $v$ . Podemos entonces calcular, para cada punto  $(x, t)$ , el valor  $a \in A$  para el cual  $\mathbf{f}(x, a) \cdot \nabla_x v(x, t) + r(x, t)$  alcanza el máximo.

En otras palabras, para cada  $(x, t)$  elegimos  $\alpha(x, t) = a$  tal que

$$\max_{a \in A} [\mathbf{f}(x, a) \cdot \nabla_x v(x, t) + r(x, t)]$$

ocurre para  $a = \alpha(x, t)$ .

Entonces, si es posible resolvemos

$$\begin{cases} d\mathbf{X}^*(s) &= \mathbf{f}(\mathbf{X}^*(s), \alpha(\mathbf{X}^*(s), s))ds + \sigma d\mathbf{W}(s) \\ \mathbf{X}^*(t) &= x. \end{cases}$$

En consecuencia,  $\mathbf{A}^*(s) = \alpha(\mathbf{X}^*(s), s)$  es el control óptimo buscado.

# Capítulo 4

## Juegos Diferenciales

En este capítulo consideraremos problemas con dos controles simultáneos. Podemos pensar que tenemos un juego entre dos personas, que llamaremos I y II, que controlan cierta situación en un intervalo de tiempo. A cada uno de los posibles resultados de la interacción le asignaremos un número. Este representará lo que el jugador II le tendrá que pagar al jugador I (de ser negativo, I estará pagándole a II). Llamaremos a dicho número el *pago del juego*. El jugador I querrá maximizarlo, y II querrá minimizarlo, ya que para él representa un *costo*.

Más formalmente, tendremos un sistema dinámico que dependerá de dos controles, uno para representar las movidas de I, que llamaremos  $\alpha(\cdot)$ , y el otro para las de II,  $\beta(\cdot)$ . Se tendrá también un funcional de pago asociado. Observemos que ahora éste dependerá de dos variables ( $\alpha(\cdot), \beta(\cdot)$ ), y que se busca maximizarlo en  $\alpha(\cdot)$ , y minimizarlo en  $\beta(\cdot)$ . O sea, queremos un punto silla ( $\alpha^*(\cdot), \beta^*(\cdot)$ ). Cuando I juegue  $\alpha^*(\cdot)$  y II  $\beta^*(\cdot)$ , diremos que estamos en *equilibrio*.

En la primera sección formalizaremos este modelo, en la Sección §2 mostraremos condiciones de optimalidad. Por último, en la sección §3 relacionaremos la noción de equilibrio con la de existencia de una solución viscosa.

### 4.1. Definiciones

Consideremos el siguiente sistema

$$\begin{cases} \dot{\mathbf{x}}(s) = f(\mathbf{x}(s), \mathbf{y}(s), \mathbf{z}(s)) & t \leq s \leq T \\ \mathbf{x}(t) = x \end{cases} \quad (4.1.1)$$

donde  $0 \leq t < T$ , y  $x \in \mathbb{R}^n$ ; dos funciones medibles  $\mathbf{y} : [t, T] \rightarrow \mathbf{Y}$  y  $\mathbf{z} : [t, T] \rightarrow \mathbf{Z}$  donde  $\mathbf{Y} \subseteq \mathbb{R}^k$ ,  $\mathbf{Z} \subseteq \mathbb{R}^l$  son conjuntos compactos.

Supongamos además que  $f : \mathbb{R}^n \times \mathbf{Y} \times \mathbf{Z} \rightarrow \mathbb{R}^n$  es uniformemente continua, y existe

una constante positiva  $C_1$  tal que

$$\begin{aligned} |f(x, y, z)| &\leq C_1 \\ |f(x_1, y, z) - f(x_2, y, z)| &\leq C_1|x_1 - x_2| \end{aligned}$$

para todo  $0 \leq t < T$ ,  $x_1, x_2 \in \mathbb{R}^n$ ,  $y \in \mathbf{Y}$ ,  $z \in \mathbf{Z}$ . Bajo estas condiciones, este sistema tiene solución única  $\mathbf{x}(\cdot)$ , que dependerá de la elección de las funciones  $\mathbf{y}(\cdot)$  y  $\mathbf{z}(\cdot)$ .

Se tiene el siguiente funcional de pago asociado al sistema

$$P(\mathbf{y}, \mathbf{z}) = P_{x,t}[\mathbf{y}(\cdot), \mathbf{z}(\cdot)] = \int_t^T r(\mathbf{x}(s), \mathbf{y}(s), \mathbf{z}(s))dt + g(\mathbf{x}(T)). \quad (4.1.2)$$

donde  $g : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  satisface

$$\begin{aligned} |g(x)| &\leq C_2 \\ |g(x_1) - g(x_2)| &\leq C_2|x_1 - x_2| \end{aligned}$$

y  $r : \mathbb{R}^n \times \mathbf{Y} \times \mathbf{Z} \rightarrow \mathbb{R}$  es uniformemente continua, tal que

$$\begin{aligned} |r(x, y, z)| &\leq C_3 \\ |r(x_1, y, z) - r(x_2, y, z)| &\leq C_3|x_1 - x_2| \end{aligned}$$

para ciertas constantes positivas fijas  $C_2$  y  $C_3$ , y para todo  $0 \leq t \leq T$ ,  $x_1, x_2 \in \mathbb{R}^n$ ,  $y \in \mathbf{Y}$ ,  $z \in \mathbf{Z}$ .

**Definición 4.1.1.** *Llamamos juego diferencial al sistema (4.1.1) junto con el pago (4.1.2) asociado.*

**Definición 4.1.2.** *Los conjuntos*

$$M(t) = \{\mathbf{y} : [t, T] \rightarrow \mathbf{Y} \mid \mathbf{y} \text{ medible}\}$$

$$N(t) = \{\mathbf{z} : [t, T] \rightarrow \mathbf{Z} \mid \mathbf{z} \text{ medible}\}$$

*son los controles admisibles para I y II respectivamente.*

Las funciones  $\mathbf{y}$  y  $\mathbf{z}$  antes mencionadas, se pueden interpretar como las acciones o movidas para cada uno de los jugadores. Definimos el pago como la magnitud que I quiere maximizar y II quiere minimizar, y que depende de la elección de estas funciones. Entonces se podría pensar que estamos maximizando sobre el conjunto  $M(t)$  y minimizando sobre  $N(t)$ . Sin embargo, esta no sería una representación correcta de la situación, pues la idea de que I maximice sobre todos sus controles admisibles no representa el hecho de que la función objetivo depende también de las acciones de II, que está minimizando en simultáneo.

Un mejor modelo será considerar a I maximizando sobre todas sus posibles *respuestas* a las acciones de II, y viceversa. Esto motiva la siguiente definición:

**Definición 4.1.3.** Una estrategia para I (empezando a tiempo  $t$ ) será cualquier asignación

$$\alpha : N(t) \rightarrow M(t)$$

tal que para cada  $s \in [t, T]$ ,  $\mathbf{z}_1, \mathbf{z}_2 \in N(t)$

$$\begin{aligned} & \text{si } \mathbf{z}_1(\tau) = \mathbf{z}_2(\tau) \text{ c.t.p para todo } \tau \in [t, s], \\ & \text{entonces } \alpha(\mathbf{z}_1)(\tau) = \alpha(\mathbf{z}_2)(\tau) \text{ para todo } \tau \in [t, s]. \end{aligned}$$

Análogamente, una estrategia para II será cualquier asignación

$$\beta : M(t) \rightarrow N(t)$$

tal que para cada  $s \in [t, T]$ ,  $\mathbf{y}_1, \mathbf{y}_2 \in M(t)$

$$\begin{aligned} & \text{si } \mathbf{y}_1(\tau) = \mathbf{y}_2(\tau) \text{ c.t.p para todo } \tau \in [t, s], \\ & \text{entonces } \beta(\mathbf{y}_1)(\tau) = \beta(\mathbf{y}_2)(\tau) \text{ para todo } \tau \in [t, s]. \end{aligned}$$

Interpretamos una estrategia para I como una función que a cada posible acción  $\mathbf{z}(\cdot)$  del jugador II, le asigna una respuesta  $\mathbf{y}(\cdot) = \alpha(\mathbf{z}(\cdot))$ .

El conjunto

$$\Gamma(t) = \{\alpha : N(t) \rightarrow M(t)\}$$

representa todas las estrategias para I, y

$$\Delta(t) = \{\beta : M(t) \rightarrow N(t)\}$$

todas las estrategias para II.

Luego, se estará maximizando el funcional de pago  $P$  sobre  $\Gamma(t)$ , y minimizando sobre  $\Delta(t)$ .

**Definición 4.1.4.** Definimos el valor superior del juego como

$$\begin{aligned} u(x, t) &= \sup_{\alpha \in \Gamma(t)} \inf_{\mathbf{z} \in N(t)} P(\alpha(\mathbf{z}), \mathbf{z}) \\ &= \sup_{\alpha \in \Gamma(t)} \inf_{\mathbf{z} \in N(t)} \left\{ \int_t^T r(\mathbf{x}(s), \alpha(\mathbf{z}(s)), \mathbf{z}(s)) ds + g(\mathbf{x}(T)) \right\}, \end{aligned}$$

donde  $\mathbf{x}(\cdot)$  es la solución de (4.1.1) para  $\mathbf{y}(\cdot) = \alpha(\mathbf{z}(\cdot))$  y  $\mathbf{z}(\cdot)$ .

Análogamente, el valor inferior del juego es

$$\begin{aligned} v(x, t) &= \inf_{\beta \in \Delta(t)} \sup_{\mathbf{y} \in M(t)} P(\mathbf{y}, \beta(\mathbf{y})) \\ &= \inf_{\beta \in \Delta(t)} \sup_{\mathbf{y} \in M(t)} \left\{ \int_t^T r(\mathbf{x}(s), \mathbf{y}(s), \beta(\mathbf{y}(s))) ds + g(\mathbf{x}(T)) \right\}, \end{aligned}$$

donde  $\mathbf{x}(\cdot)$  es la solución de (4.1.1) para  $\mathbf{y}(\cdot)$  y  $\mathbf{z}(\cdot) = \beta(\mathbf{y}(\cdot))$ .

*Observación 4.1.1.* Si consideramos la situación en que los jugadores se turnan para mover, ir primero es una desventaja: el que elige segundo sabe qué control eligió el anterior. La función valor inferior  $v$  representa una versión infinitesimal de esta situación, en la que el jugador  $I$  tiene ventaja. En este caso,  $I$  tiene una estrategia que le asegura ganar *al menos*  $v$ , de aquí el nombre de valor inferior.

Análogamente,  $u$  representa la situación en que  $II$  tiene ventaja, y, como éste está minimizando su costo, le da una estrategia que le asegura perder *a lo sumo*  $u$ .

Luego, tenemos  $v(x, t) \leq u(x, t)$ .

## 4.2. Propiedades del Valor Superior e Inferior

Los siguientes teoremas caracterizan al valor superior y al inferior. Sus demostraciones serán omitidas (pueden verse en [5]) pues son muy similares a las demostraciones del Principio de Programación Dinámica 2.2.1 y el Teorema 2.2.2 que vimos en el caso de un problema de control.

**Teorema 4.2.1.** *Para cada  $h > 0$ , suficientemente chico tal que  $t + h \leq T$ , se tiene*

$$u(x, t) = \sup_{\alpha \in \Gamma(t)} \inf_{\mathbf{z} \in N(t)} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \alpha(\mathbf{z}(s)), \mathbf{z}(s)) ds + u(t+h, \mathbf{x}(t+h)) \right\},$$

$$v(x, t) = \inf_{\beta \in \Delta(t)} \sup_{\mathbf{y} \in M(t)} \left\{ \int_t^{t+h} r(\mathbf{x}(s), \mathbf{y}(s), \beta(\mathbf{y}(s))) ds + v(t+h, \mathbf{x}(t+h)) \right\}.$$

**Teorema 4.2.2.** *Si  $u, v \in C^1$ , entonces  $u$  es una solución viscosa de la ecuación superior de Isaacs,*

$$u_t + \min_{b \in B} \max_{a \in A} \{ \mathbf{f}(x, a, b) \cdot \nabla_x u(x, t) + r(x, a, b) \} = 0$$

$$u(x, T) = g(x)$$

*y  $v$  es una solución viscosa de la ecuación inferior de Isaacs,*

$$v_t + \max_{a \in A} \min_{b \in B} \{ \mathbf{f}(x, a, b) \cdot \nabla_x v(x, t) + r(x, a, b) \} = 0$$

$$v(x, T) = g(x)$$

Las ecuaciones de Isaacs para un juego de suma cero con dos jugadores son análogas a las de Hamilton-Jacobi-Bellman. Podemos reescribirlas como

$$u_t + H^+(x, \nabla_x u) = 0$$

para el Hamiltoniano superior

$$H^+(x, p) := \min_{b \in B} \max_{a \in A} \{ \mathbf{f}(x, a, b) \cdot p + r(x, a, b) \}$$

y

$$v_t + H^-(x, \nabla_x v) = 0$$

para el Hamiltoniano inferior

$$\max_{a \in A} \min_{b \in B} \{ \mathbf{f}(x, a, b) \cdot p + r(x, a, b) \}$$

*Observaciones 4.2.1.* En las siguientes observaciones comparamos el teorema minimax de Von Neumann, donde el conjunto de estrategias de cada jugador es finito, y el pago del juego está dado por una matriz, con nuestra situación.

i) En general, se tiene

$$\max_{a \in A} \min_{b \in B} \{ \mathbf{f}(x, a, b) \cdot p + r(x, a, b) \} < \min_{b \in B} \max_{a \in A} \{ \mathbf{f}(x, a, b) \cdot p + r(x, a, b) \}$$

y, en consecuencia,

$$H^-(x, p) < H^+(x, p).$$

La demostración es la misma que en el caso discreto (ver, por ejemplo, [1]), pero en el caso continuo las ecuaciones superior e inferior de Isaacs resultan diferentes, y entonces los valores superior e inferior del juego también lo son:  $u \neq v$ .

ii) Si para todo  $x, p$  se tiene

$$\max_{a \in A} \min_{b \in B} \{ \mathbf{f}(x, a, b) \cdot p + r(x, a, b) \} = \min_{b \in B} \max_{a \in A} \{ \mathbf{f}(x, a, b) \cdot p + r(x, a, b) \},$$

decimos que el juego satisface las condiciones *minimax*, también llamadas de *Isaacs*. En este caso,  $u \equiv v$ , y decimos que el juego tiene un *valor*. En esta situación, podríamos diseñar controles óptimos para los jugadores *I* y *II*.

iii) Decir que  $\alpha^*(\cdot)$  y  $\beta^*(\cdot)$  son óptimos, quiere decir que el par  $(\alpha^*(\cdot), \beta^*(\cdot))$  es un *punto silla* de  $P_{x,t}$ , es decir

$$P_{x,t}[\alpha(\cdot), \beta^*(\cdot)] \leq P_{x,t}[\alpha^*(\cdot), \beta^*(\cdot)] \leq P_{x,t}[\alpha^*(\cdot), \beta(\cdot)] \quad (4.2.1)$$

para todo par de controles  $\alpha(\cdot), \beta(\cdot)$ .

iv) El jugador *I* elige  $\alpha^*(\cdot)$  si supone que *II* jugará  $\beta^*(\cdot)$  porque es su mejor respuesta, y, a su vez, *II* elige  $\beta^*(\cdot)$ , porque si cree que *I* jugará  $\alpha^*(\cdot)$ , también es su mejor respuesta. En Teoría de Juegos, un punto  $(\alpha^*(\cdot), \beta^*(\cdot))$  donde cada estrategia es una mejor respuesta a la estrategia elegida por el rival se llama un *Equilibrio de Nash*.

v) En un juego arbitrario pueden existir numerosos equilibrios de Nash, cada uno con diferentes pagos, y cada jugador elegirá una estrategia haciendo suposiciones sobre cuál será el equilibrio que el otro jugador estará intentando jugar. Eventualmente, pueden elegir estrategias correspondientes a equilibrios de Nash distintos.

- vi) En general, en los juegos de suma cero la desigualdad (4.2.1) nos garantiza que hay un pago único al que ambos jugadores pueden aspirar, el *valor del juego*, pero podrían existir diferentes pares de estrategias óptimas  $(\alpha_i^*(\cdot), \beta_i^*(\cdot))$ . No es difícil verificar que se tiene

$$P_{x,t}[\alpha_i^*(\cdot), \beta_i^*(\cdot)] = P_{x,t}[\alpha_j^*(\cdot), \beta_j^*(\cdot)] = P_{x,t}[\alpha_i^*(\cdot), \beta_j^*(\cdot)]$$

para todo  $i, j$ . Sin embargo, bajo la condición de Isaacs, Evans y Souganidis demostraron que existe un único par de estrategias óptimas que son las soluciones viscosas del Teorema 4.2.2.

Para finalizar, observemos que uno podría estar interesado en estudiar un problema con  $N$  jugadores. Sin embargo, aparecen diferentes complicaciones que dificultan el uso de este enfoque, pues por ejemplo ya no hay un valor inferior y uno superior, sino  $N!$  órdenes posibles para jugar.

Avner Friedman propuso en [6] un enfoque para estos problemas. La idea es fijar una permutación  $\pi$  del orden en que juegan los jugadores, subdividir el intervalo  $[0, T]$  en intervalos de longitud  $\delta$ , y en cada uno, cada jugador elige su estrategia consecutivamente conociendo las estrategias elegidas por todos en todos los subintervalos anteriores, y las estrategias elegidas por los jugadores que lo preceden en el subintervalo en cuestión. Luego, se hace tender  $\delta \rightarrow 0$ , obteniéndose un valor  $V_\pi$ . Si es independiente de la permutación, se puede definir el valor del juego.

Sin embargo, esto no cubre todos los casos posibles, pues debe decidirse si los jugadores pueden o no formar coaliciones, asociándose para jugar en conjunto. En caso de estar permitidas estas coaliciones, se debe analizar si los pagos son o no transferibles entre los jugadores de una coalición, y resolver el problema del reparto del pago entre los jugadores.

En el próximo capítulo veremos un nuevo enfoque propuesto por Lasry y Lions en 2007, los juegos de campo medio.

# Capítulo 5

## Juegos de campo medio

En este capítulo veremos dos ejemplos de juegos de campo medio. En el primero no interesará la posición de los jugadores, solamente el instante que eligen para llegar a una reunión. El segundo, más completo, analiza el problema de elegir un nivel de tecnología para aislar una vivienda tratando de minimizar los costos de inversión, de mantenimiento, y de energía para calefaccionarla. En este último incluimos la deducción del sistema de ecuaciones propuestas por Lasry y Lions para modelar los juegos de campo medio.

En cada caso, la existencia de solución dependerá de distintas técnicas de ecuaciones diferenciales: el primero se resuelve utilizando un teorema de punto fijo, el segundo se transforma en una ecuación ordinaria que se integra explícitamente, y para el tercero se debe hallar el minimizante de un funcional.

### 5.1. ¿Cuándo empieza una reunión?

Consideremos el problema del horario real en que empieza una reunión. Es muy común que, aunque ésta se programe para empezar cierto horario  $t$ , comience unos minutos más tarde. Este tiempo, que llamaremos  $T$ , depende de la dinámica de llegada de los participantes. Si se tiene alguna regla para determinar cuándo comenzará la reunión, se generará una interacción estratégica entre los concurrentes.

Cada agente  $i$  elige un tiempo  $\tau^i$  al que quiere llegar, acorde a cierto costo (o pago) que quiere minimizar (maximizar). Es decir, cada participante *controla* su tiempo de llegada, y busca optimizar esta decisión.

Si bien cada concurrente decide en qué momento  $\tau^i$  le gustaría llegar, en realidad llegará en el instante  $\tilde{\tau}^i = \tau^i + \sigma^i \varepsilon^i$ , donde  $\varepsilon^i$  es un ruido normal con varianza 1 (los supondremos independientes), y la variante específica para cada agente  $i$  está dada por  $\sigma_i > 0$ . Más precisamente,  $\tau^i$  es la variable controlada por  $i$ , y  $\sigma^i \varepsilon^i$  es una incertidumbre a la que cada agente está sujeto. Estas incertidumbres y sus intensidades difieren en la población, representan factores externos, como, por ejemplo, el tráfico, o las distintas

distancias que recorre cada uno para llegar. Notaremos con  $m_0$  la distribución de  $\sigma^i$  en la población.

Recordemos entonces los elementos de este modelo:

- $t$  es el tiempo (horario) al que está programada la reunión.
- $\tau^i$  es el tiempo en que le gustaría llegar a cada agente  $i$ .
- $\tilde{\tau}^i = \tau^i + \sigma^i \varepsilon^i$  será el tiempo en que realmente llega, donde  $\varepsilon^i$  es un ruido normal con varianza 1, y  $\sigma^i$  denotará una incertidumbre para cada agente. Supondremos que  $\sigma^i > c$ ,  $\forall i$  para alguna constante  $c > 0$ .
- $T$  es el tiempo en que la reunión efectivamente comenzará.
- $m_0$  es la distribución de  $\sigma^i$  en la población.

Para decidir en qué horario quiere llegar, cada agente optimizará un costo total. En el caso más simple, consideraremos la función costo

$$c(t, T, \tilde{\tau}) = c_1(t, T, \tilde{\tau}) + c_2(t, T, \tilde{\tau}) + c_3(t, T, \tilde{\tau})$$

donde, para ciertas constantes  $\alpha$ ,  $\beta$  y  $\gamma$  positivas,

$$c_1(t, T, \tilde{\tau}) = \alpha(\tilde{\tau} - t)^+$$

es el costo (efecto en la reputación) de llegar tarde en relación al tiempo acordado  $t$ ,

$$c_2(t, T, \tilde{\tau}) = \beta(\tilde{\tau} - T)^+$$

es el costo (inconveniente personal) de llegar tarde en relación al tiempo de inicio real  $T$ , y

$$c_3(t, T, \tilde{\tau}) = \gamma(T - \tilde{\tau})^+$$

es el costo (tiempo de espera) del tiempo que se pierde esperando hasta  $T$ .

Observemos que  $c(t, T, \tilde{\tau})$  es una función convexa de  $\tilde{\tau}$ .

Cada agente quiere minimizar el costo total esperado. Según la teoría de Equilibrios de Nash con expectativas racionales, sobre la cual no profundizaremos, cada agente optimiza asumiendo  $T$  como conocido.  $T$  es *a priori* una variable aleatoria, pero como consideramos un número infinito de jugadores, la Ley de los Grandes Números implica que  $T$  es determinística, y así la consideraremos en adelante.

Para cada agente  $i$  el problema será hallar

$$\tau^i = \operatorname{argmin} E[c((t, T, \tilde{\tau}))].$$

Notemos que acá  $T$  es el campo medio, la estimación exhaustiva para cada agente del comportamiento de los otros.

Nuestro objetivo es mostrar que existe un punto fijo  $T$ , esto es, probar que las optimizaciones individuales, suponiendo conocido  $T$ , general la realización de este tiempo  $T$ , que llamaremos *equilibrio*.

Para mostrar que dicho equilibrio existe, primero examinemos las decisiones individuales.

**Proposición 5.1.1.** *El tiempo optimo  $\tau^i$  de un agente con incertidumbre  $\sigma^i > c > 0$  está definido implícitamente por*

$$\alpha\Phi\left(\frac{\tau^i - t}{\sigma^i}\right) + (\beta + \gamma)\Phi\left(\frac{\tau^i - T}{\sigma^i}\right) = \gamma$$

donde  $\Phi$  es la función de distribución acumulada de una variable aleatoria normal.

*Demostración.* La expresión que se quiere minimizar es

$$\mathbb{E}[\alpha(\tilde{\tau}^i - t)^+ + \beta(\tilde{\tau}^i - T)^+ + \gamma(T - \tilde{\tau}^i)^+],$$

que podemos escribir como

$$\mathbb{E}[\alpha(\tilde{\tau}^i - t)^+ + (\beta + \gamma)(\tilde{\tau}^i - T)^+ - \gamma(\tilde{\tau}^i - T)],$$

y por la linealidad de la esperanza, queda

$$\alpha\mathbb{E}[(\tau^i - t + \sigma^i\varepsilon^i)^+] + (\beta + \gamma)\mathbb{E}[(\tau^i - T + \sigma^i\varepsilon^i)^+] - \gamma(\tau^i - T).$$

Derivando en  $\tau_i$ , tenemos

$$\begin{aligned} \frac{d}{d\tau^i}\mathbb{E}[(\tau^i - t + \sigma^i\varepsilon^i)^+] &= \frac{d}{d\tau^i} \int_{\{\tau^i - t + \sigma^i\varepsilon^i > 0\}} (\tau^i - t + \sigma^i\varepsilon^i) dZ = \\ &= \int_{\{\tau^i - t + \sigma^i\varepsilon^i > 0\}} dZ = \mathbb{P}(\tau^i - t + \sigma^i\varepsilon^i > 0), \end{aligned}$$

donde  $dZ$  denota integrar respecto a la distribución Normal. Por lo tanto, tenemos

$$\alpha\mathbb{P}(\tau^i - t + \sigma^i\varepsilon^i > 0) + (\beta + \gamma)\mathbb{P}(\tau^i - T + \sigma^i\varepsilon^i > 0) = \gamma.$$

Si llamamos  $\Phi$  a la función de distribución de la Normal, es decir, si  $W \sim \mathcal{N}(0, 1)$  entonces

$$\mathbb{P}(W \leq w) = \Phi(w),$$

y además, por la simetría de la normal respecto del 0, vale que

$$\Phi(w) = \mathbb{P}(W \leq w) = \mathbb{P}(W \geq -w)$$

Entonces, la condición anterior nos queda

$$\alpha\Phi\left(\frac{\tau^i - t}{\sigma^i}\right) + (\beta + \gamma)\Phi\left(\frac{\tau^i - T}{\sigma^i}\right) = \gamma. \quad (5.1.1)$$

Por ser  $\Phi$  una función de distribución estrictamente monótona, y  $\alpha$ ,  $\beta$  y  $\gamma$  positivas, se deduce la existencia y unicidad de  $\tau^i$ .  $\square$

De esta caracterización de  $\tau^i$  como función de  $(t, T, \sigma^i)$ , podemos deducir la dinámica de llegada de los agentes. Para esto, consideremos la distribución  $m_0$  de  $\sigma^i$ . Notaremos por  $F$  a la función de distribución acumulada de los tiempos reales de llegada de los agentes. Como suponemos un continuo de agentes, y por la Ley de los Grandes Números,  $F$  puede ser considerada determinística.

Necesitamos establecer una regla para dar comienzo a la reunión, es decir, que determine el tiempo  $T$ . Es natural pensar que dicha regla dependerá de  $F(\cdot)$ . Supondremos, por ejemplo, que se rige por *quorum*: la reunión comienza luego del horario programado y sólo cuando una proporción  $\theta$  de las participantes han llegado.

Vamos a probar ahora la existencia y unicidad de un punto fijo. Comenzando de un valor  $T$ , obtenemos las estrategias óptimas de los agentes  $(\tau^i(\cdot; T))_i$ . Si bien estos son los óptimos de cada agente, sabemos que el tiempo real de llegada de cada uno de ellos está afectado por un ruido; necesitamos obtener entonces  $(\tilde{\tau}(\cdot; T))_i$ . Luego, por la Ley de los Grandes Números y la hipótesis de independencia de las incertidumbres de los agentes, estos tiempos de llegada tienen la distribución  $F$ . Recordemos, de nuevo, que  $F$  es determinística, y  $T$  se deduce de  $F$  con la regla de inicio,  $(T^*(F))$ , en este caso el quorum. El siguiente esquema resume lo anterior:

$$T^{**} : T \mapsto (\tau^i(\cdot; T))_i \mapsto (\tilde{\tau}(\cdot; T))_i \mapsto F = F(\cdot; T) \mapsto T^*(F)$$

*Observación 5.1.1.* La regla de quorum  $T^*(F)$  satisface:

- La reunión nunca empieza antes de  $t$ ,

$$\forall F(\cdot), T^*(F(\cdot)) \geq t,$$

- (Monotonía) Sean  $F(\cdot), G(\cdot)$  dos funciones de distribución acumulada,

$$F(\cdot) \leq G(\cdot) \Rightarrow T^*(F(\cdot)) \geq T^*(G(\cdot)).$$

- (Sub-aditividad)

$$\forall s > 0, T^*(F(\cdot - s)) - T^*(F(\cdot)) \leq s.$$

Esta condición es la menos natural de las tres. Consideramos  $G(\cdot) = F(\cdot - s)$ , es decir, la situación en la que todos los agentes llegan  $s$  instantes más tarde que en el caso  $F(\cdot)$ , pero la distribución es la misma. Entonces, es razonable esperar que haya quorum a más tardar  $s$  instantes después que cuando hay quorum para la distribución original, es decir,  $T^*(F(\cdot - s)) \leq s + T^*(F(\cdot))$ . Esta idea es similar a la del Principio de Programación Dinámica: dejamos pasar  $s$  instantes y empezamos a contar con  $F(\cdot)$ .

Con esto, obtenemos el siguiente resultado:

**Proposición 5.1.2.** Si  $\alpha > 0$ ,  $\beta > 0$ ,  $\gamma > 0$  y  $0 \notin \text{Sop}(m_0)$ , entonces  $T^{**}$  es una contracción como función de  $[t, +\infty[$ , y existe una única solución  $T$  para nuestro problema.

*Demostración.* Primero derivemos con respecto a  $T$  la condición de primer orden (5.1.1) que define a  $\tau^i$ :

$$\begin{aligned}
0 &= \frac{\partial}{\partial T} \left( \alpha \Phi\left(\frac{\tau_i - t}{\sigma_i}\right) + (\beta + \gamma) \Phi\left(\frac{\tau_i - T}{\sigma_i}\right) \right) \\
&= \alpha \Phi'\left(\frac{\tau_i - t}{\sigma_i}\right) \frac{\partial\left(\frac{\tau_i - t}{\sigma_i}\right)}{\partial T} + (\beta + \gamma) \Phi'\left(\frac{\tau_i - T}{\sigma_i}\right) \frac{\partial\left(\frac{\tau_i - T}{\sigma_i}\right)}{\partial T} \\
&= \alpha \Phi'\left(\frac{\tau_i - t}{\sigma_i}\right) \frac{\partial}{\partial \tau_i} \left( \frac{\tau_i - t}{\sigma_i} \right) \frac{\partial \tau_i}{\partial T} \\
&\quad + (\beta + \gamma) \left[ \Phi'\left(\frac{\tau_i - T}{\sigma_i}\right) \left( \frac{\partial}{\partial \tau_i} \left( \frac{\tau_i - T}{\sigma_i} \right) \frac{\partial \tau_i}{\partial T} \left( \frac{-1}{\sigma_i} \right) \right) \right] \\
&= \alpha \Phi'\left(\frac{\tau_i - t}{\sigma_i}\right) \frac{1}{\sigma_i} \frac{\partial \tau_i}{\partial T} - \frac{(\beta + \gamma)}{\sigma_i} \Phi'\left(\frac{\tau_i - T}{\sigma_i}\right) \\
&\quad + (\beta + \gamma) \Phi'\left(\frac{\tau_i - T}{\sigma_i}\right) \frac{\partial \tau_i}{\partial T} \frac{1}{\sigma_i}.
\end{aligned}$$

Despejando nos queda

$$\frac{d\tau^i}{dT} \left[ \alpha \Phi'\left(\frac{\tau^i - t}{\sigma^i}\right) + (\beta + \gamma) \Phi'\left(\frac{\tau^i - T}{\sigma^i}\right) \right] = (\beta + \gamma) \Phi'\left(\frac{\tau^i - T}{\sigma^i}\right),$$

con lo cual,

$$\frac{d\tau^i}{dT} \leq \frac{(\beta + \gamma) \Phi'\left(\frac{\tau^i - T}{\sigma^i}\right)}{\alpha \Phi'\left(\frac{\tau^i - t}{\sigma^i}\right) + (\beta + \gamma) \Phi'\left(\frac{\tau^i - T}{\sigma^i}\right)} \leq k \leq 1,$$

pero como 0 no está en el soporte de  $m_0$ , tenemos que  $\frac{d}{dT} \tau(t, \sigma, T) \leq k < 1$ .

Entonces, para todo  $T, s, h > 0$ ,

$$\begin{aligned}
F(s, T + h) &= \mathbb{P}(\tau^i(\sigma^i, T + h) + \sigma^i \varepsilon^i \leq s) \\
&\geq \mathbb{P}(\tau^i(\sigma^i, T) + kh + \sigma^i \varepsilon^i \leq s) \\
&= F(s - kh, T)
\end{aligned}$$

En consecuencia

$$T^*(F(\cdot, T + h)) \leq T^*(F(\cdot - kh, T)) \leq T^*(F(\cdot, T)) + kh,$$

de donde concluimos

$$T^{**}(T + h) - T^{**}(T) \leq kh,$$

es decir,  $T^{**}$  es una contracción, y entonces existe  $T$  punto fijo.  $\square$

*Observación 5.1.2.* Cuando planteamos el modelo, pedimos  $\sigma^i > 0$ . Esto implica, en particular,  $0 \notin \text{Sop}(m_0)$ .

Es interesante observar que el quorum no es un caso especial, en el sentido de que esta demostración sigue valiendo para cualquier regla de inicio  $\tilde{T} : F(\cdot) \mapsto T$  que verifique las propiedades enunciadas en (5.1.1).

Es natural pensar que a los agentes no sólo les importará el tiempo real de inicio de la reunión  $T$ , sino también la dinámica de llegada de los demás agentes. Es decir, el costo no sólo dependerá de  $T$ , sino también de  $F$ . En este caso, más general,  $F$  es el mean field: cada persona toma su decisión basándose en  $F$ , y a su vez,  $F$  se construye sobre esas decisiones. Desde el punto de vista matemático, el punto fijo depende de  $F$ .

## 5.2. El problema de la calefacción

### 5.2.1. El modelo

Supongamos que se tiene un número de jugadores  $N$  muy grande, cada uno con una casa y los recursos suficientes para acondicionarla y calefaccionarla. No se retirarán del juego hasta un cierto tiempo  $T > 0$ .

Los jugadores intentan minimizar sus costos. Si bien toman las decisiones en forma individual, se ven influenciados por las decisiones que tomaron los restantes jugadores (sólo en los costos, no hay efectos manada, ni moda).

Cada jugador quiere estar en su casa a una temperatura agradable, y para eso debe decidir si invierte en un revestimiento aislante, o si paga un costo mayor en electricidad (energía) para calefaccionarla.

Elegir cómo va a revestir su casa es fijar un  $x \in \mathbb{R}$ : si  $x \searrow -\infty$ , pierde toda la aislación; si  $x \nearrow +\infty$ , la casa adquiere toda la tecnología disponible.

### 5.2.2. Distribución de los jugadores

La aislación a tiempo  $t$ ,  $X_t$ , varía de la siguiente forma:

$$dX_t = \sigma dW + \alpha(t, X_t)dt$$

- El primer término es un movimiento browniano, y representa los factores climáticos que pueden acelerar o frenar el deterioro del aislante.
- $\sigma$  es una constante dada, relacionada con las innovaciones en tecnología, y el clima (calidad de los materiales, intensidad de las tormentas, etc.)
- El término  $\alpha(t, x)$  es la inversión para aumentar (o disminuir) el revestimiento de la casa.

Suponiendo que el número de jugadores tiende a infinito, denotamos por  $m_0$  la densidad inicial de los jugadores en la recta. Esta describe la proporción de jugadores con un nivel de aislamiento dado a tiempo  $t = 0$ . Llamaremos  $m(t, \cdot)$  a la densidad de los agentes, para cada  $t$  en  $[0, T]$ . La dinámica de  $m(t, \cdot)$  estará dada por:

$$m_t - \frac{\sigma^2}{2}m'' + (\alpha m)' = 0 \quad (x, t) \in \mathbb{R} \times [0, T]$$

Esta es la ecuación de Kolmogorov para el proceso de difusión. Veamos su deducción.

**Proposición 5.2.1** (Ecuación de Kolmogorov (Forward)). *Si  $X_t$  es un proceso estocástico de la forma*

$$dX_t = b dW + a(t, X_t) dt,$$

*con una distribución inicial  $m_0$ , la evolución de la distribución de  $X_t$  está dada por*

$$m_t - \frac{b^2}{2}m'' + (am)' = 0$$

*Demostración.* Primero vamos a derivar la ecuación de Kolmogorov para el caso  $a(t, X_t)$  constante. Para eso, vamos a utilizar paseos al azar en tiempo discreto. Esto es, partimos un intervalo de tiempo, de longitud  $t$ , en  $n = t/\Delta t$  “escalones” discretos, y en cada escalón, con probabilidad  $p$   $X_t$  aumenta en  $\Delta h$ , y con probabilidad  $q = 1 - p$ , disminuye  $\Delta h$ , siendo

$$p = \frac{1}{2}\left[1 + \frac{a}{b}\sqrt{\Delta t}\right], \quad q = \frac{1}{2}\left[1 - \frac{a}{b}\sqrt{\Delta t}\right]$$

Queremos, además, mantener la esperanza y la varianza de  $X_t - m_0$  independiente de la elección de  $\Delta t$ . Tomamos entonces

$$\Delta h = b\sqrt{\Delta t}$$

Sea  $m(\cdot, t)$  la densidad de probabilidad de  $X_t$ , dada la distribución inicial  $m(x, 0) = m_0(x)$ . Entonces

$$\mathbb{P}(c \leq X_t \leq d) = \int_c^d m(u, t) du.$$

En el intervalo  $t - \Delta t$  a  $t$ , el proceso puede alcanzar el punto  $x$  de dos formas, o bien creciendo desde  $x - \Delta h$ , o bien decreciendo de  $x + \Delta h$ . Luego

$$m(x, t) = pm(x - \Delta h, t - \Delta t) + qm(x + \Delta h, t - \Delta t). \quad (5.2.1)$$

Desarrollemos  $m(x - \Delta h, t - \Delta t)$  en serie de Taylor alrededor de  $m(x, t)$ :

$$m(x - \Delta h, t - \Delta t) = m(x, t) - \Delta t m_t - \Delta h m_x + \frac{1}{2}(\Delta h)^2 m_{xx} + \dots$$

Observemos que los términos de tercer orden y superiores son de orden  $(\Delta t)^{\frac{3}{2}}$ ,  $(\Delta t)^2$ , etc. y, por lo tanto, se van a cero más rápido que  $\Delta t$ . Ahora hagamos el desarrollo de Taylor alrededor de  $m(x, t)$  para  $m(x + \Delta h, t - \Delta t)$ :

$$m(x + \Delta h, t - \Delta t) = m(x, t) - \Delta t m_t + \Delta h m_x + \frac{1}{2}(\Delta h)^2 m_{xx} + \dots$$

Suponemos que podemos descartar los términos de mayor orden, igual que en la demostración de la regla de Ito, y sustituímos todo en (5.2.1)

$$\begin{aligned} m(x, t) &= p[m(x, t) - \Delta t m_t - \Delta h m_x + \frac{1}{2}(\Delta h)^2 m_{xx}] \\ &\quad + q[m(x, t) - \Delta t m_t + \Delta h m_x + \frac{1}{2}(\Delta h)^2 m_{xx}] = \\ &= (p + q)m(x, t) - (p + q)\Delta t m_t - (p - q)\Delta h m_x + \frac{1}{2}(\Delta h)^2 m_{xx}. \end{aligned}$$

Usemos que  $p + q = 1$ , y, por como definimos  $p$  y  $q$ ,  $p - q = \frac{a}{b}\sqrt{\Delta t}$ . Recordemos además que  $\Delta h = b\sqrt{\Delta t}$ :

$$\Delta t m_t + \left(\frac{a}{b}\right)\sqrt{\Delta t}\Delta h m_x - \frac{1}{2}(\Delta h)^2 m_{xx} = 0$$

Reemplacemos  $\Delta h = b\sqrt{\Delta t}$ :

$$\Delta t m_t + a\Delta t m_x - \frac{1}{2}b^2\Delta t m_{xx} = 0$$

Ahora dividamos todo por  $\Delta t$ :

$$m_t + a m_x - \frac{b^2}{2} m_{xx} = 0$$

que es la ecuación Forward de Kolmogorov. De forma similar, podemos derivar esta ecuación para el caso  $a(X_t, t)$  no necesariamente constante.  $\square$

### 5.2.3. Costos

Los jugadores quieren minimizar un costo, dado por tres componentes:

- El primer término es  $\alpha^2(t, x)$ , la inversión en aislar la casa.
- Otro término es de la forma  $f(t, x) = p(t)[1 - \beta(x)]$  que representa el costo de la electricidad para calefaccionar la casa;  $\beta : \mathbb{R} \rightarrow [0, 1]$  es una función no decreciente que dice cuánto se ahorra en energía por tener revestimiento.

- El último término es

$$g(t, x, m) = \frac{a_1 \gamma(x)}{1 + a_2 m(t, x)}, \quad a_1, a_2 > 0.$$

Aquí,  $\gamma : \mathbb{R} \rightarrow [0, 1]$  es monótona creciente,  $\lim_{x \rightarrow -\infty} \gamma(x) = 0$ . Este término es creciente en  $x$ , porque a mayor inversión en revestimiento, mayor es el gasto de mantenimiento; pero es decreciente en  $m(t, x)$  porque el costo se abarata por razones de economía de escala al ser muchos los que tienen el mismo nivel o tecnología de aislamiento  $x$ .

#### 5.2.4. El problema

Queremos hallar el control  $\alpha(t)$  y la distribución  $m$  que minimicen

$$\inf_{\alpha} \mathbb{E} \left[ \int_0^T \alpha^2(t, X_t) + f(t, X_t) + g(t, X_t, m(t, X_t)) dt \right]$$

con la dinámica para  $X_t$  dada por

$$dX_t = \sigma dW + \alpha(t, X_t) dt,$$

y una distribución inicial  $m_0(x)$ .

Usando la ecuación Forward de Kolmogorov, tenemos que la distribución de jugadores resuelve

$$\begin{aligned} m_t - \frac{\sigma^2}{2} m'' + (\alpha m)' &= 0 & (x, t) \in \mathbb{R} \times [0, T], \\ m(x, 0) &= m_0(x) & x \in \mathbb{R}. \end{aligned}$$

Por el Capítulo 3, sabemos que el problema de control estocástico puede reducirse a una ecuación de Hamilton-Jacobi-Bellman,

$$v_t + \frac{\sigma^2}{2} v_{xx} + \max_a [av_x + a^2 + g(t, x, m(t, x))] = -f(t, x),$$

e imponemos una condición terminal

$$v(\cdot, T) = 0,$$

pues no hay un costo terminal.

En general, el sistema es difícil de resolver, y si calculamos la esperanza a minimizar,

$$\inf_{\alpha} \left\{ \int_{\mathbb{R}} \int_0^T [\alpha^2(t, x) + f(t, x) + g(t, x, m(t, x))] m(t, x) dt dx \right\},$$

en este caso podemos probar que existe una distribución de jugadores  $m(t, x)$  y un control  $\alpha(t, x)$  que minimizan el costo con herramientas del cálculo de variaciones.

### 5.2.5. Existencia de un ínfimo

Por comodidad, tomamos  $\sigma = 1$ , y definimos

$$q(t, x) = \alpha(t, x)m(t, x),$$

con lo cual, para recuperar  $\alpha$ , introducimos

$$\psi(a, b) = \begin{cases} \frac{|a|^2}{b} & \text{si } b > 0 \\ +\infty & \text{caso contrario.} \end{cases}$$

Llamemos

$$\varphi(m)(t) = \int_{\mathbb{R}} [f(t, x) + g(t, x, m(t, x))]m(t, x)dx,$$

que está bien definida pues  $m$  es integrable, y  $f, g$  son acotadas. Definimos también

$$K(q, m) = \int_0^T \int_{\mathbb{R}} \psi(q, m)dxdt + \int_0^T \varphi(m)(t)dt,$$

y elegimos el siguiente espacio para minimizar  $K$

$$B := \left\{ (q, m) : q \in L^2([0, T] \times \mathbb{R}), m \text{ solución débil de } m_t - m'' = -q', \right. \\ \left. m(0, \cdot) = m_0(\cdot) \in L^2(\mathbb{R}), m \in L^2([0, T], H^1) \right\}$$

donde  $' = \frac{d}{dx}$ ,  $'' = \frac{d^2}{dx^2}$ . Luego, podemos reescribir el problema de minimización como

$$\inf_{(q, m) \in B} K(q, m) \tag{5.2.2}$$

La existencia de una solución de (5.2.2) es un problema abierto. Sin embargo, podemos estudiar el problema penalizado. Dado  $\varepsilon > 0$ , definimos

$$\inf_{(q, m) \in B} K_\varepsilon(q, m) := K(q, m) + \varepsilon \|q\|_2^2 \tag{5.2.3}$$

donde  $\|\cdot\|_2$  denota la norma  $L^2([0, T] \times \mathbb{R})$ . El siguiente teorema prueba que, para cada  $\varepsilon > 0$ , existe un par  $(q_\varepsilon, m_\varepsilon) \in B$  minimizante de (5.2.3), y que el límite cuando  $\varepsilon$  tiende a cero es un ínfimo de (5.2.2).

**Teorema 5.2.1.** *Si  $m_0 \in H^1(\mathbb{R})$  y  $g$  es continua, entonces el problema de minimización (5.2.3) admite una solución  $(q_\varepsilon, m_\varepsilon) \in B$ , con  $q_\varepsilon \in L^2([0, T] \times \mathbb{R})$ . Más aún,*

$$\lim_{\varepsilon \rightarrow 0} \left( \min_{(q, m) \in B} K_\varepsilon(q, m) \right) = \inf_{(q, m) \in B} K(q, m)$$

Antes de la demostración, vamos a enunciar el siguiente teorema, cuya demostración puede encontrarse en [13].

**Teorema 5.2.2.** *Sea  $f \in L^2([0, T], H^{-1})$ ,  $m_0 \in H^1$ , entonces existe una solución débil  $m$  de*

$$m_t - m'' = f$$

que satisface

$$m \in L^2([0, T], H^1) \cap L^\infty([0, T], H^1).$$

Ahora sí podemos demostrar el Teorema 5.2.1.

*Demostración del Teorema 5.2.1.* Dado  $\varepsilon > 0$  fijo, existe  $(q_n, m_n)$  una sucesión minimizante del problema (5.2.3), por ser  $K_\varepsilon(q, m) \geq 0$ . Vamos a organizar la demostración en pasos, y cuando sea necesario tomar una subsucesión, la seguiremos notando  $\{q_n\}_n$  ó  $\{m_n\}_n$ :

1. **Cotas y convergencia de  $q_n$ .** Existe una constante  $c > 0$  tal que

$$\varepsilon \|q_n\|_2^2 \leq K_\varepsilon(q_n, m_n) \leq c.$$

Luego, la sucesión  $q_n$  es acotada uniformemente en  $L^2([0, T] \times \mathbb{R})$ , y, en consecuencia, existe una subsucesión (que volveremos a llamar  $q_n$ ) que converge débil a  $q \in L^2([0, T] \times \mathbb{R})$ .

Observemos además, que  $q'_n$  están acotadas uniformemente en  $L^2([0, T], H^{-1})$ .

2. **Convergencia de  $m_n$ .** Como  $q'_n \in L^2([0, T], H^{-1})$  y  $m_0 \in H^1(\mathbb{R})$ , por el teorema (5.2.2), tenemos que  $m_n \in L^2([0, T], H^1)$ . Más aún, como  $q'_n$  está acotada en  $L^2([0, T], H^{-1})$ ,  $m_n$  está acotada en  $L^2([0, T], H^1)$ .

Queremos ver que podemos extraer una subsucesión de  $m_n$  que converge fuerte a un límite  $m \in L^2([0, T] \times \mathbb{R})$ . Se tienen las siguientes inclusiones continuas:

$$H^1 \subset L^2(\mathbb{R}) \subset H^{-1},$$

donde la primera inclusión es compacta por el Teorema de Inmersión. Temam demostró que esto implica que la inclusión  $L^2([0, T], H^1) \subset L^2([0, T] \times \mathbb{R})$  también es compacta, una demostración de esta afirmación puede encontrarse en [13].

Luego, como  $m_n$  están acotadas uniformemente, existe  $m \in L^2([0, T], H^1)$ , y una subsucesión  $m_n$  tal que  $m_n$  converge débilmente a  $m$  en  $L^2([0, T], H^1)$  y fuertemente en  $L^2([0, T] \times \mathbb{R})$ .

Por último, podemos extraer otra subsucesión, de nuevo la llamamos  $\{m_n\}_n$  que converge a  $m$  en casi todo punto.

Observemos que, como  $q_n \rightharpoonup q$  en  $L^2([0, T] \times \mathbb{R})$  y  $m_n \rightharpoonup m$  en  $L^2([0, T], H^1)$ , entonces  $m$  es solución débil de la ecuación de Komogorov, y por lo tanto,  $(q, m) \in B$ . Esto se verifica utilizando el núcleo del calor, y pasando al límite. Tenemos

$$m_n(t, x) = G(t) \star m_0(x) - \int_0^t G(t-s) \star \frac{d}{dx} q_n(s) ds,$$

y, en sentido débil,

$$G(t-s) \star \frac{d}{dx} q_n(s) = \frac{d}{dx} G(t-s) \star q_n(s) \in C^\infty(\mathbb{R}),$$

con lo cual podemos pasar al límite en la sucesión  $m_n$ :

$$\begin{aligned} \lim_{n \rightarrow \infty} m_n(t, x) &= \lim_{n \rightarrow \infty} \left[ G(t) \star m_0(x) - \int_0^t \frac{d}{dx} G(t-s) \star q_n(s, x) ds \right] \\ &= G(t) \star m_0(x) - \int_0^t \frac{d}{dx} G(t-s) \star q(s, x) ds \\ &= m(t, x). \end{aligned}$$

3. **Paso al límite.** Por la semicontinuidad de la norma  $L^2([0, T] \times \mathbb{R})$  y del término de la izquierda en  $K$ , por la convergencia en casi todo punto de  $m_n$  a  $m$ , la continuidad de  $g$  y el lema de Fatou, tenemos:

$$\begin{aligned} K_\varepsilon(q, m) &= \lim_{n \rightarrow \infty} K_\varepsilon(q_n, m_n) \\ &= \liminf_{n \rightarrow \infty} K_\varepsilon(q_n, m_n) \\ &\geq \int_0^T \int_{\mathbb{R}} \frac{|q|^2}{m} + \int_0^T \varphi(m) + \varepsilon \|q\|_2^2 \\ &= K_\varepsilon(q, m), \end{aligned}$$

lo que muestra que  $(q, m)$  es un minimizante de (5.2.3).

4. **Límite de los problemas penalizados.** Sea  $(q_\varepsilon, m_\varepsilon)$  una solución de (5.2.3), entonces, para todo par  $(q, m) \in B$ :

$$\inf_{(q, m) \in B} K(q, m) \leq K(q_\varepsilon, m_\varepsilon) + \varepsilon \|q_\varepsilon\|_2^2 \leq K(q, m) + \varepsilon \|q\|_2^2.$$

Tomemos ahora el límite de  $\varepsilon$  tendiendo a 0, :

$$\begin{aligned} \inf_{(q,m) \in B} K(q, m) &\leq \liminf_{\varepsilon \rightarrow 0} K_\varepsilon(q_\varepsilon, m_\varepsilon) \\ &\leq \limsup_{\varepsilon \rightarrow 0} K_\varepsilon(q_\varepsilon, m_\varepsilon) \\ &\leq K(q, m), \end{aligned}$$

Tomemos ahora ínfimo sobre los pares  $(q, m) \in B$ :

$$\begin{aligned} \inf_{(q,m) \in B} K(q, m) &\leq \liminf_{\varepsilon \rightarrow 0} K_\varepsilon(q_\varepsilon, m_\varepsilon) \\ &\leq \limsup_{\varepsilon \rightarrow 0} K_\varepsilon(q_\varepsilon, m_\varepsilon) \\ &\leq \inf_{(q,m) \in B} K(q, m), \end{aligned}$$

que es lo que queríamos probar.

□

*Observación 5.2.1.* La existencia de una solución de la ecuación (5.2.2) es un problema abierto. Si se pudiese probar que:

$$\exists M > 0 \text{ tal que } \forall \varepsilon > 0, \|q_\varepsilon\|_2 < M,$$

donde  $(q_\varepsilon, m_\varepsilon) \in B$  es una solución del problema penalizado (5.2.3), entonces el problema (5.2.2) admite una solución  $(q, m) \in B$ , con  $q \in L^2([0, T] \times \mathbb{R})$ .



# Bibliografía

- [1] K. Binmore, *Playing for Real*, Oxford University Press, USA, 2007.
- [2] L.C. Evans, *Partial Differential Equations*. American Math Society, 1998.
- [3] L.C. Evans, *An Introduction to Mathematical Optimal Control Theory*. Lecture notes: Version 0.2.
- [4] L. C. Evans, *An Introduction to Stochastic Differential Equations*. Lecture notes: Version 1.2.
- [5] . L.C. Evans and P.E. Souganidis, *Differential games and representation formulas for solutions of Hamilton-Jacobi equations*. Indiana University Mathematics Journal 33 (1984), 773-797.
- [6] A. Friedman, *Linear quadratic differential games with nonzero sum and with  $N$  players*. Archive Rat. Mech., Anal., Vol. 34 (1969), 165-187.
- [7] O. Gueant, *Théorie des jeux à champ moyen et applications économiques. Taux d'escompte et développement durable*. Tesis doctoral, Université Paris-Dauphine, junio de 2009.
- [8] O. Gueant, J.-M. Lasry and P.-L. Lions, *Mean field games and applications*, en *Paris-Princeton Lectures on Mathematical Finance 2010*, ed. A. Cousin, S. Crépey, O. Guéant, y D. Hobson. Lecture Notes in Mathematics, Springer, 2010.
- [9] A. Lachapelle, *Quelques problèmes de transport et de contrôle en économie: aspects théoriques et numériques*. Tesis doctoral, Université Paris-Dauphine, junio de 2010.
- [10] J.-M. Lasry y P.-L. Lions, *Jeux à champ moyen. I. Le cas stationnaire*. C. R. Math. Acad. Sci. Paris, vol. 343 (2006) 619-625.
- [11] J.-M. Lasry y P.-L. Lions, *Jeux à champ moyen. II. Horizon fini et contrôle optimal*. C. R. Math. Acad. Sci. Paris , vol. 343 (2006) 679-684.

- [12] J.-M. Lasry y P.-L. Lions, *Mean field games*, Jpn. J. Math., vol. 2 (2007) 229-260.
- [13] R. Temam, *Navier-Stokes Equations and Nonlinear Functional Analysis*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol 41.