Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional

Unidad Zacatenco

Departamento de Matemáticas

Control óptimo de interacción de partículas en tiempo discreto

Tesis que presenta

Israel Sanchez Mendoza

Para obtener el grado de

MAESTRO EN CIENCIAS

EN LA ESPECIALIDAD DE MATEMÁTICAS

Asesor de Tesis:

Dr. Héctor Jasso Fuentes, CINVESTAV-IPN, México

Ciudad de México Febrero, 2021

Índice general

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) ya que gracias a la beca propocionada se logró la realización de este proyecto (#CVU 782866).

Agradezco el apoyo del Centro de Investigación y Estudios Avanzados del IPN (CINVESTAV), en particular a su Departamento de Matemáticas por el apoyo recibido durante mis estudios, ya que sin ellos este trabajo no sería posible. así como el tiempo invertido por todos mis profesores.

Agradezco también profundamente a mi director de tesis Héctor Jasso Fuentes por haberme brindado la oportunidad de trabajar en conjunto y recurrir a su capacidad y conocimientos, así como por el apoyo y paciencia incondicional durante mis estudios y durante el desarrollo de este proyecto.

Asimismo agradezco el trabajo y esfuerzo de los sinodales en revisar este trabajo de tesis así como las sugerencias recibidas.

Adicionalmente agradezco el apoyo incondicional de mi familia, ya que gracias a ellos me encuentro en este punto de mi vida, y que todo se los debo a ellos. En especial este trabajo es en tu memoria Víctor.

Por último quiero dedicar esta tesis a Berenice ya que me ha apoyado en cada decisión que he tomado y me ha brindado el apoyo y ánimo necesario para la conclusión de este proyecto, así como por todo lo dado.

Abstract

This thesis is about the study of a discrete-time optimal control problem composed of large number of interacting particles. In order to do so, a general model of N-particles is first analyzed; within this, the associated control problem to be optimized is posed as well as its solvability through the use of the dynamic programming technique. From the theoretical point of view, the existence of optimal policies will be shown; however, in practice, obtaining such optimal policies could be very complex, due to the size of N.

To avoid this drawback, we make use of mean-field theory; i.e., we set up a "limit" control model when $N \to \infty$, called the mean-field model. This limit model is also studied rigorously in this work, obtaining in particular its optimal elements, such as optimal costs and optimal policies. The latter are much simpler to obtain, as they do not depend on N.

An important result is to prove the convergence of the N-particles model to the mean-field model. This fact is achieved by approximating their paths and their costs. A consequence of this, is to be able of measuring the performance of optimal mean-field policies in the N-particles model when N is sufficiently large, which is very useful in applications.

Finally, in order to illustrate the theory, two applications are presented in the work. The first one has to do with a consumption-investment model, in which N investors interact and we aim for an optimal policy that allows minimizing the government or regulator's subdsidy/tax costs. The second application is concerned with the allocation of requests in a computer network, in which we wish to minimize the time it takes to attend those requests.

Resumen

Esta tesis trata el estudio de un problema de control óptimo de interacción de partículas en tiempo discreto. Para tal fin, primero se analiza un modelo general de N-partículas y se plantea su problema de control a optimizar, así como la solubilidad del mismo a través del uso de la técnica de programación dinámica. En términos teóricos, se verá y garantizará la existencia de políticas óptimas; sin embargo, en la práctica, el mostrar explícitamente la política óptima, puede ser muy complejo, debido al tamaño de N.

Para evitar este inconveniente, se hace uso de la teoría de campo medio; es decir, se plantea un modelo de control "límite" cuando $N \to \infty$, denominado modelo de campo medio. Este modelo límite también se estudia de manera rigurosa en este trabajo, obteniéndose en particular sus elementos óptimos, como lo son los costos óptimos y políticas óptimas. Estos últimos son mucho más sencillos de obtenerse, debido a que no dependen de N.

Un resultado importante es el demostrar la convergencia del modelo de N partículas al modelo de campo medio. Esto se hace a través de la aproximación de sus trayectorias y de sus costos. Una consecuencia de esto es el poder medir el rendimiento de políticas óptimas de campo medio en el modelo de N partículas cuando N es suficientemente grande, lo cual es muy útil en aplicaciones.

Finalmente, a fin de ilustrar la teoría, se establecen dos aplicaciones. La primera tiene que ver con un modelo de consumo-inversión dentro del cual interactúan N inversionistas y buscamos una política óptima que permita minimizar los costos de subdsidio/impuesto del gobierno o de la entidad regulatoria. La segunda aplicación tiene que ver con la asignación de solicitudes en una red computacional, en el que se desea minimizar el tiempo de atención de las solicitudes.

1

Introducción

Los problemas de control óptimo (PCO) han sido estudiados a lo largo de varias décadas debido a su utilidad en diversas áreas como la ingeniería, la economía, la biología, las matemáticas financieras, entre otras. Una razón de su estudio es que las circunstancias del mundo actual, como lo son: la escacez de recursos, la sobre-población, etc., han orillado a que muchos de los problemas de hoy en día se deban de resolver de manera óptima para ahorrar recursos de diversa índole. Es ahí donde la teoría de control óptimo tiene gran relevancia.

Recordemos que un PCO puede ser visto como un problema de optimización en donde se aplican reglas o decisiones a través del tiempo con la finalidad de optimizar recursos limitados. Este tipo de problemas puede ser clasificados desde diversos ángulos dependiendo del tipo de aplicacion a tratar, ya que sus modelos asociados pueden ser de tipo determinístico o estocástico; o pueden ser a tiempo discreto o a tiempo contínuo; o el tipo de función objetivo que se desee optimizar puede ser de un tipo específico, etc... En cada uno de estos casos, existe una gran cantidad de trabajos teóricos y prácticos que han dado soluciones desde el punto de vista analítico/teórico y numérico/práctico.

En años recientes ha existido la necesidad de estudiar sistemas que involucran una cantidad extremadamente grande de poblaciones, mismas que pueden representar partículas, objetos, seres vivos, etc. En ocasiones también se da la necesidad de controlar este tipo de problemas, pero resulta que el número de partículas involucradas en el modelo, digamos N, está muy ligada con su solubilidad; en particular, resulta que la dimensión del estado del sistema, es la misma que N. Sin embargo, existe una teoría que puede darle solución a estos problemas, la cual consiste en encontrar un modelo de control "límite" de tal manera que el modelo de N partículas "converge" a dicho modelo límite cuando $N \to \infty$. La idea detrás de esta convergencia, es reescribir el modelo de N partículas en término de sus proporciones y posteriormente tomar el límite cuando N se va a infinito.

En esta tesis trabajamos el modelo estocástico de control "centralizado" de interacción de partículas en tiempo discreto, cuando se tiene un número *finito* de clases (o estados) dentro de las (los) cuales cada partícula toma valores. La función objetivo a analizar es un costo total esperado con horizonte de planeación finito. El término de control centralizado significa que, en cada etapa del tiempo, un controlador central observa el estado de la posición/distribución de las partículas y una vez logrado eso, el controlador toma una acción que se encuentra dentro de un conjunto dado. Como consecuencia de lo anterior, se generan dos situaciones: (1) un costo por “actuar” y (2) una transición (o movimiento) aleatorio de todas las partículas dentro de las diferentes clases.

La finalidad del trabajo es mostrar la eficiencia del modelo de campo medio, asegurando la existencia de políticas óptimas y costos óptimos sobre el modelo de campo medio, con la finalidad de comparar estos objetos óptimos con el modelo de interés de N- partículas y así poder de alguna manera “resolver” o aproximar los objetos óptimos de este último mencionado. Los detalles de cómo se realizarán tales acciones se establecen en la Sección 1.2.

1.1 Literatura relacionada

Los problemas de controlabilidad de campo medio tienen sus orígenes en los trabajos de Lasry y Lions [15] y de forma independiente de Huang [13, 12], concernientes ambos a juegos de campo medio. A partir de estos trabajos, este tipo de modelos se han multiplicado de manera considerable tanto en la parte teórica como en aplicaciones en economía, física estadística, investigación de operaciones, entre otros —ver, por ejemplo, [1, 2, 3, 6, 7, 14, 16, 17, 18] y las referencias contenidas en estas. Cabe mencionar que esta tesis se ha inspirado en los trabajos de control de campo medio plasmados en las referencias [5, 9, 10]. Nuestra finalidad es mostrar un material de autocontenido sobre parte del análisis visto en estos artículos, añadiendo a su vez algunos ejemplos que sustentan la teoría.

1.2 Contenido del trabajo

El contenido de este trabajo se describe como sigue. Después de esta sección introductoria, se presenta el modelo de control de N partículas dentro del cual se define el estado del sistema, las acciones y controles que puede aplicar el controlador central y las dinámicas tanto del movimiento de cada partícula y como del ambiente que las rodea. Posteriormente se verá la manera de cambiar el monitoreo de cada partícula (monitoreo microscópico) a un monitoreo de las proporciones de las mismas (monitoreo macroscópico). También se define la función objetivo y el problema de control asociado al modelo de N partículas. El capítulo concluye con la solución teórica aplicando la teoría de programación dinámica. Sin embargo, se menciona que en términos prácticos el problema se vuelve muy complejo de resolver debido a que la cardinalidad N del número de partículas influye en la solubilidad práctica.

Por otro lado, el Capítulo 3 trata de la definición del modelo de campo medio, el cual es determinístico y no depende de N. Se introducen todos sus elementos (similares al modelo de N partiículas), incluyendo el problema de optimización a resolver. De igual manera, se prueba la optimalidad de este modelo utilizando la programación dinámica y se concluye que la solubilidad de este problema es sencilla de obtener.

Con el material de los dos capítulos anteriores, se da pauta a los resultados de convergencia, mismos que se analizan en el Capítulo 4. Específicamente se supone que las trayectorias del problema de N partículas convergen en cierto sentido a las trayectorias del modelo de campo medio. Posteriormente se hace un análisis de convergencia tanto de los costos por etapa, como la convergencia de las funciones objetivo; en particular para los costos óptimos entre ambos modelos. Esto último trae como consecuencia el poder utilizar políticas de control óptimas del modelo de campo medio que pueden ser utilizadas como "buenos aproximadores" de políticas óptimas para el modelo de N partículas, lo que constituye una simplificación y facilidad en la práctica.

La tesis finaliza con el Capítulo 5 que tiene que ver con el desarrollo de algunas aplicaciones, en donde se pueden apreciar las utilidades del modelo de control de campo medio. Primeramente nos centramos en un modelo de consumo-inversión, dentro del cual analizamos el planteamiento del problema y el análisis consistirá en demostrar de manera precisa y rigurosa que las hipótesis A y B que se plantean en la teoría, se satisfacen de manera natural. Esto da como resultado el poder asegurar la existencia de políticas óptimas en el modelo de campo medio que sirvan como buenos aproximadores del modelo original. Adicionalmente, se analiza de manera más resumida un modelo de asignación de tareas en una red computacional, el cual formularemos como un modelo de interacción de partículas; sin embargo, la manera de comprobar que las hipótesis antes mencionadas se satisfacen, es muy parecida a la del primer ejemplo, por lo que solamente haremos puntualizaciones del modelo.

2

Modelo de control markoviano de N-partículas

En todo este trabajo supondremos la existencia de un espacio de probabilidad fijo $(\Omega, \mathcal{F}, \mathbb{P})$, sobre el cual se definirán una serie de elementos descritos más adelante.

2.1 Sistema con N partículas

Considere un sistema compuesto por N partículas u objetos, donde cada partícula puede ser medida mediante su *estado o clase* del sistema, la cual vive en el conjunto $\mathcal{S} := \{1, 2, 3, \ldots, S\}$.

Para cada $1 \leq n \leq N$ denotaremos por $X_n^N(t) \in \mathcal{S}$ el estado de la n-ésima partícula en el tiempo $t \in \mathbb{N}_0$. De esta forma, el estado completo del sistema compuesto por las N partículas en cada etapa de tiempo t se define como:

$$\mathtt{X}^N(t) := \left(X_1^N(t), X_2^N(t), \ldots, X_N^N(t)\right) \in \mathcal{S}^N.$$

Además asumiremos la existencia de un agente externo (o controlador) el cual, mediante la elección de una acción $\mathtt{a}(t) \in A$ (siendo A un espacio métrico), tiene la posibilidad de controlar el sistema en cada etapa del tiempo t. Finalmente, vamos a asumir que las partículas comparten un ambiente en común, que también influirá en el comportamiento del sistema. A dicho ambiente en el tiempo t lo denotaremos por $\mathtt{C}^N(t) \in \mathbb{R}^d$.

A partir de los elementos anteriores, la ecuación de estado para la partícula n-ésima del proceso está definida por la siguiente dinámica:

$$X_n^N(t+1) = F\left(X_n^N(t), \mathtt{C}^N(t), \mathtt{a}(t), \xi_t^n\right), \quad X_n^N(0) = x_0^N, \ t \geq 0, \tag{2.1}$$

donde $F : \mathcal{S} \times \mathbb{R}^d \times A \times \mathbb{R} \to \mathcal{S}$ es una función medible, $\{\xi_t^n : t \in \mathbb{N}, n = 1, \ldots, N\}$ es una familia de variables aleatorias i.i.d. con valores en $\mathbb{R}$ y con distribución común ρ definida en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.

Así que, para cada acción $\mathtt{a}(t) \in A$ y contexto del ambiente $\mathtt{C}^N(t) \in \mathbb{R}^d$, tenemos que la probabilidad de transición de que la partícula n-ésima se mueva del estado i al estado j, estará dada por:

$$\begin{aligned} K_{ij}^{\rho}\left(\mathtt{a},\mathtt{c}\right) &:= \mathbb{P}\left(X_n^N\left(t+1\right)=j|X_n^N\left(t\right)=i,\mathtt{a}\left(t\right)=\mathtt{a},\mathtt{c}^N\left(t\right)=\mathtt{c}\right) \\ &= \int_{\mathbb{R}} \mathbb{1}_j\left(F\left(i,\mathtt{c},\mathtt{a},z\right)\right)\rho\left(dz\right). \end{aligned} \tag{2.2}$$

Se puede observar que dichas probabilidades de transición son idénticas para cualquier partícula n, de esta manera podemos decir que, las partículas se mueven independientemente unas de otras.

Es importante señalar que cuando N es muy grande, el controlador no sería capaz de seguir el comportamiento de cada partícula, por el contrario, se asumirá que éste únicamente puede observar el número de partículas en cada clase del sistema, i.e. en cada elemento de $\mathcal{S}$.

Así pues, el sistema se observará de manera macroscópica, sondeando en cada etapa de tiempo, el número de partículas en cada estado $i \in \mathcal{S}$. Con lo cual la proporción de dichas partículas en el estado $i \in \mathcal{S}$ al tiempo t se encuentra dada por:

$$M_i^N\left(t\right) := \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{\{X_n^N(t)=i\}}. \tag{2.3}$$

Esto conlleva a definir el estado (macroscópico) de las proporciones del sistema conforme a lo siguiente:

$$\mathtt{M}^N\left(t\right) = \left(M_1^N\left(t\right), M_2^N\left(t\right), \ldots, M_S^N\left(t\right)\right), \tag{2.4}$$

donde podemos ver que

$$\mathtt{M}^N\left(t\right) \in \mathcal{P}_N\left(\mathcal{S}\right) := \left\{p \in \mathcal{P}\left(\mathcal{S}\right) : Np\left(i\right) \in \mathbb{N}, \forall i \in \mathcal{S}\right\},$$

donde $\mathcal{P}\left(\mathcal{S}\right)$ es el conjunto de medidas de probabilidad sobre $\mathcal{S}$.

Ahora bien, con este nuevo enfoque, la ecuación que determinará la evolución del contexto del ambiente es:

$$\mathtt{C}^N\left(t+1\right) = g\left(\mathtt{C}^N\left(t\right), \mathtt{M}^N\left(t+1\right), \mathtt{a}\left(t\right)\right), \tag{2.5}$$

donde $g : \mathbb{R}^d \times \mathcal{P}_N\left(\mathcal{S}\right) \times A \to \mathbb{R}^d$ es una función medible.

Finalmente suponemos que la probabilidad de transición K (en específico, la probabilidad ρ) afecta la evolución de $\mathtt{M}^N\left(t\right)$ de manera recursiva, i.e., asumiremos que existe una función medible $G_\rho^N : \mathcal{P}_N\left(\mathcal{S}\right) \times \mathbb{R}^d \times A \times \mathbb{R}^N \to \mathcal{P}_N\left(\mathcal{S}\right)$ tal que

$$\mathtt{M}^N\left(t+1\right) = G_\rho^N\left(\mathtt{M}^N\left(t\right), \mathtt{C}^N\left(t\right), \mathtt{a}\left(t\right), \mathtt{w}_t\right), \tag{2.6}$$

donde $\left(\mathtt{w}_t\right)_{t\in\mathbb{N}}$ es una sucesión de vectores aleatorios i.i.d. en $\mathbb{R}^N$ con distribución común θ definida también en el espacio de probabilidad $\left(\Omega, \mathcal{F}, \mathbb{P}\right)$. Tal consideración no es nada restrictiva; de hecho, la representación en (2.6) se establece por medio de técnicas para simular cadenas de Markov. La relación entre las proporciones dadas por medio de (2.3) y por medio (2.6) no es trivial y se encuentran fuera del alcance de esta tesis. Sin embargo, el lector puede consultar el Teorema 3.3 en [2] —véase también [5, 9, 10].

Considere ahora el espacio producto $\mathcal{Y}_N := \mathcal{P}_N(\mathcal{S}) \times \mathbb{R}^d$. Y definimos la función $H_\rho^N : \mathcal{Y}_N \times A \times \mathbb{R}^N \to \mathcal{Y}_N$ dada por

$$H_\rho^N(\mathtt{y}, a, \mathtt{w}) := \left(G_\rho^N(\mathtt{y}, a, \mathtt{w}), g\left(\mathtt{c}, G_\rho^N(\mathtt{y}, \mathtt{a}, \mathtt{w}), a\right)\right) \quad \forall\, (\mathtt{y}, a, \mathtt{w}) \in \mathcal{Y}_N \times A \times \mathbb{R}^N. \tag{2.7}$$

Esta función define la dinámica del proceso $\mathtt{Y}^N(t) := \left(\mathtt{M}^N(t), \mathtt{C}^N(t)\right)$, de la siguiente manera:

$$\begin{aligned} \mathtt{Y}^N(t+1) &= \left(G_\rho^N\left(\mathtt{Y}^N(t), \mathtt{a}(t), \mathtt{w}_t\right), g\left(\mathtt{C}^N(t), \mathtt{M}^N(t+1), \mathtt{a}(t)\right)\right) \\ &= \left(G_\rho^N\left(\mathtt{Y}^N(t), \mathtt{a}(t), \mathtt{w}_t\right), g\left(\mathtt{C}^N(t), G_\rho^N\left(\mathtt{Y}^N(t), \mathtt{a}(t), \mathtt{w}_t\right), \mathtt{a}(t)\right)\right) \\ &= H_\rho^N\left(\mathtt{Y}^N(t), \mathtt{a}(t), \mathtt{w}_t\right). \end{aligned} \tag{2.8}$$

Por otro lado, definimos de igual manera el espacio como $\mathcal{Y} := \mathcal{P}(\mathcal{S}) \times \mathbb{R}^d$. Asumiendo así que existe una función medible

$$\ell : \mathcal{Y} \times A \to \mathbb{R}, \tag{2.9}$$

que representará el costo en el que incurre el agente por la decisión (o acción) que toma en cada una de las etapas. Dicho costo dependerá del estado del sistema $\mathtt{Y}^N(\cdot) = \left(\mathtt{M}^N(\cdot), \mathtt{C}^N(\cdot)\right)$, así como de la acción seleccionada por el agente.

De forma similiar supondremos que existe otra función medible:

$$\ell_T : \mathcal{Y} \to \mathbb{R}, \tag{2.10}$$

la cual representará el costo terminal incurrido al llegar al estado final. Cabe señalar, que dicho costo dependerá exclusivamente del estado del sistema.

2.2 N-Modelo de control

De acuerdo a los conceptos y definiciones que usamos en el modelo de interacción de N partículas, aplicaremos la teoría de control óptimo para minimizar cierto costo total, el cual se definirá más adelante.

Para empezar, formularemos el N-Modelo de Control Markoviano (N-MCM) al sistema de N-partículas como la séxtupla:

$$\mathcal{M}_N := \left(\mathcal{Y}_N, A, H_\rho^N, \theta, \ell, \ell_T\right), \tag{2.11}$$

cuyos elementos han sido definidos en la sección anterior.

Estos elementos determinarán el sistema estocástico $\mathtt{Y}^N(t) = \left(\mathtt{M}^N(t), \mathtt{C}^N(t)\right)$ en el que evolucionan los estados a través del tiempo, los cuales son las proporciones de las partículas en los estados y el ambiente en el que interactúan.

Cabe mencionar que si el sistema al tiempo $t \in \{0, \cdots, T-1\}$ se encuentra en $\mathtt{Y}^N(t) = \mathtt{y} \in \mathcal{Y}_N$ y se toma la acción $\mathtt{a}(t) = a \in A$ entonces sucede lo siguiente:

1. Se genera el costo $\ell\left(\mathtt{y},a\right)$ de forma instantánea.
2. El sistema se mueve a un nuevo estado $\mathtt{Y}^N\left(t+1\right)=\left(\mathtt{M}^N\left(t+1\right),\mathtt{C}^N\left(t+1\right)\right)$ de acuerdo a la probabilidad de transición

$$\begin{aligned} Q_\rho\left(B|\mathtt{y},a\right) &= \mathbb{P}\left(\mathtt{Y}^N\left(t+1\right)\in B|\mathtt{Y}^N\left(t\right)=\mathtt{y},\mathtt{a}\left(t\right)=a\right) \\ &= \int_{\mathbb{R}^N}\mathbb{1}_B\left(H_\rho^N\left(\mathtt{y},a,w\right)\right)\theta\left(dw\right) \\ &= \mathbb{E}\left[\mathbb{1}_B\left(H_\rho^N\left(\mathtt{y},a,\mathtt{w}\right)\right)\right]. \end{aligned} \tag{2.12}$$

3. Si $t=T$ y el estado del sistema se encuentra en $\mathtt{Y}(T)=\mathtt{y}\in\mathcal{Y}_N$, entonces se generará el costo $\ell_T\left(\mathtt{y}\right)$ y el sistema se detiene.

Ahora bien, las acciones seleccionadas por el controlador son hechas de acuerdo a reglas que llamaremos *políticas de control*, las cuales se definirán a continuación.

Sea $\mathcal{H}_0^N:=\mathcal{Y}_N$. De forma recursiva definimos el espacio de historias al tiempo t como $\mathcal{H}_t^N:=\left(\mathcal{Y}_N\times A\right)^t\times\mathcal{Y}_N$ para $t=1,2,\cdots T$.

Cabe destacar que un elemento de $\mathcal{H}_t^N$ está dado conforme a lo siguiente:

$$h_t^N=\left(\mathtt{y}_0^N,a_0,\ldots,\mathtt{y}_{t-1}^N,a_{t-1},\mathtt{y}_t^N\right)\in\mathcal{H}_t^N\quad\forall\ t=1,2,\cdots T.$$

Definición 2.2.1. (a) *Una política de control es una sucesión $\pi^N=\left\{\pi_t^N\right\}_{t\in\mathbb{N}}$ de núcleos estocásticos π_t^N en A dado $\mathcal{H}_t^N$ tales que $\pi_t^N\left(A|h_t^N\right)=1$ para cada $h_t^N\in\mathcal{H}_t^N, t=0,\cdots,T-1$. Al conjunto de todas las políticas de control lo denotamos por Π^N.*

(b) *Denotaremos por $\mathbb{F}$ al conjunto de todas las funciones medibles $f:\mathcal{Y}\to A$ y $\mathbb{F}|_{\mathcal{Y}_N}$ la restricción de $\mathbb{F}$ en $\mathcal{Y}_N$. Diremos que una política de control $\pi^N\in\Pi^N$ es una política determinística si para cada $t\in\mathbb{N}_0$ existe $f_t^N\in\mathbb{F}|_{\mathcal{Y}_N}$ tal que $\pi_t^N(\cdot|h_t^N)=\delta_{f_t^N(\mathtt{y}_t^N)}(\cdot)$, $\forall\ t\in\mathbb{N}_0$.*

Escribiremos por $\mathbb{F}^N$ al conjunto de políticas determinísticas. Análogamente y abusando un poco de la notación, identificaremos al conjunto de políticas determinísticas sobre $\mathcal{Y}$ por medio del conjunto $\mathbb{F}$.

2.3 El proceso estado-acción

Definimos al proceso de historias de la siguiente manera:

$$\mathtt{h}_t^N=\left(\mathtt{Y}^N\left(0\right),\mathtt{a}\left(0\right),\ldots,\mathtt{Y}^N\left(t-1\right),\mathtt{a}\left(t-1\right),\mathtt{Y}^N\left(t\right)\right)\in\mathcal{H}_t^N. \tag{2.13}$$

Dada la condición inicial $\mathtt{y}\in\mathbb{Y}_N$ y una política $\pi^N\in\Pi^N$, introducimos el proceso *estado - acción* $\left\{\left(\mathtt{Y}^N\left(t\right),\mathtt{a}\left(t\right)\right)\right\}_{t\geq 0}$ de acuerdo a lo siguiente:

$$\mathtt{Y}^N\left(0\right):=\mathtt{y}\quad\mathtt{a}\left(0\right)\sim\pi_0^N\left(\cdot|\mathtt{y}\right), \tag{2.14}$$

y para $t \geq 1$

$$\mathtt{Y}^N(t+1) = H_\rho^N\left(\mathtt{Y}^N(t), \mathtt{a}(t), \mathtt{w}_t\right) \qquad \mathtt{a}(t+1) \sim \pi_{t+1}^N\left(\cdot|\mathtt{h}_t^N\right), \tag{2.15}$$

con $\mathtt{h}_t^N$ como en (2.13). En la expresión de arriba, hemos usado la notación $\sim$ que quiere decir “se distribuye”.

Es claro que $\left\{\mathtt{Y}^N(t)\right\}_{t\geq 0}$ es un proceso definido en $(\Omega, \mathcal{F}, \mathbb{P})$ si se usan políticas en $\mathbb{F}^N$; la razón es que al sustituir alguna $f \in \mathbb{F}^N$ en (2.14)-(2.15) y al utilizar la medibilidad de f y de H_ρ^N, el resultado se sigue de inmediato.

En general, si se usa una política general $\pi^n \in \Pi^N$, entonces el Teorema de Ionescu-Tulcea afirma la existencia de una única medida de probabilidad $P_{\mathtt{y}}^{\pi^N}$ sobre este espacio $(\Omega, \mathcal{F})$ tal que $\forall B \in \mathcal{B}(\mathcal{Y}_N), C \in \mathcal{B}(A), D \in \mathcal{B}\left(\mathbb{R}^N\right), h_t \in \mathcal{H}_t^N$ y para cada $t \geq 0$ satisface:

(1) $P_{\mathtt{y}}^{\pi^N}\left(\mathtt{Y}^N(0) \in B\right) = \delta_{\mathtt{y}}(B)$,

(2) $P_{\mathtt{y}}^{\pi^N}\left(\mathtt{a}(t) \in C|h_t^N\right) = \pi_t^N\left(C|h_t^N\right)$,

(3) $P_{\mathtt{y}}^{\pi^N}\left(\mathtt{w}_t \in D|h_t^N\right) = \mathbb{P}(\mathtt{w}_t \in D) = \theta(D)$,

(4) $P_{\mathtt{y}}^{\pi^N}\left(\mathtt{Y}^N(t+1) \in B|h_t^N, \mathtt{a}_t\right) = Q_\rho\left(B|\mathtt{y}_t^N, \mathtt{a}_t\right) = \theta\left(w \in \mathbb{R}^N | H_\rho^N\left(\mathtt{y}_t^N, \mathtt{a}_t, w\right) \in B\right)$.

2.4 Planteamiento del problema de control

Una vez mencionadas las políticas de control a usar, así como el proceso estado-acción, podemos definir el costo total esperado de una política π^N. Dicho costo puede tener diferente naturaleza, de acuerdo al tipo de problema o aplicación que se este tratando; en particular éste puede clasificarse de acuerdo al tipo de horizonte de planeación, el cual puede ser finito o infinito.

En ambos casos el controlador busca minimizar (resp. maximizar) cierto costo (resp. recompensa) a lo largo de los tiempos $t \leq T$, donde $T \leq \infty$ es mejor conocido como el *horizonte de planeación*, mismo que puede ser determinístico o estocástico, según corresponda el caso concreto a tratar.

Para el caso de horizonte finito, el costo (función objetivo) a optimizar por el controlador es el siguiente:

$$V^N\left(\pi^N, \mathtt{y}\right) := E_{\mathtt{y}}^{\pi^N}\left[\sum_{t=0}^{T-1} \ell\left(\mathtt{Y}^N(t), \mathtt{a}(t)\right) + \ell_T\left(\mathtt{Y}^N(T)\right)\right], \tag{2.16}$$

donde la esperanza se considera con respecto a la medida de probabilidad $P_{\mathtt{y}}^{\pi^N}$ inducida por la política π^N y el estado inicial $\mathtt{Y}^N(0) = \mathtt{y}$.

Decimos que $\pi_*^N \in \Pi^N$ es una política óptima para el N-MCM (2.11) si y sólo sí satisface

$$V^N\left(\pi_*^N, \mathtt{y}\right) = \inf_{\pi^N \in \Pi^N} V^N\left(\pi^N, \mathtt{y}\right) =: V_*^N(\mathtt{y}) \qquad \forall\ \mathtt{y} \in \mathcal{Y}_N, \tag{2.17}$$

A la función V_*^N se le conoce como función de valor o costo óptimo. Así que, el problema de optimización asociado al N-MCM es determinar una política π_*^N que satisfaga lo anterior

Cabe mencionar que debido al cambio de valor de las funciones ℓ y ℓ_T a través del tiempo, comúnmente se introduce un factor de descuento, es decir, un escalar $0 < \alpha < 1$, que mide el valor en el tiempo 0 de una unidad recibida en el tiempo $0 \leq t \leq T$. Con esta reformulación la función de costo en el caso de horizonte finito tendrá la forma.

$$V^N\left(\pi^N, \mathtt{y}\right) := E_{\mathtt{y}}^{\pi^N}\left[\sum_{t=1}^{T-1} \alpha^t \ell\left(\mathtt{Y}^N(t), \mathtt{a}(t)\right) + \alpha^T \ell_T\left(\mathtt{Y}^N\right)\right], \tag{2.18}$$

la cual por supuesto se puede reducir a un costo del tipo (2.16) haciendo $\hat{\ell} = \ell \cdot \alpha^{\cdot}$ y $\hat{\ell}_T = \ell_T \cdot \alpha^T$.

Finalmente hace falta imponer ciertas condiciones de compacidad y continuidad al módelo $\mathcal{M}_N$ con el fin de que exista una política óptima definida como arriba, mismos que se establecerán en la próxima sección.

2.5 Optimalidad del N-MCM

Para garantizar la existencia del valores óptimos al problema de control óptimo descrito en la pasada sección, impondremos algunas condiciones de continuidad y compacidad a nuestro modelo.

Para lo que sigue definiremos las siguientes normas:

$$\begin{aligned} \|m\|_1 &:= \max\{|m_1|, \ldots, |m_N|\} \quad \forall\ m \in \mathcal{P}_N(\mathcal{S}), \\ \|c\|_2 &:= \max\{|c_1|, \ldots, |c_d|\} \quad \forall c \in \mathbb{R}^d, \\ \|\mathtt{y}\| &:= \max\{\|m\|_1\ ,\ \|c\|_2\} \quad \forall\ \mathtt{y} := (m, c) \in \mathcal{Y}_N. \end{aligned} \tag{2.19}$$

Y de forma similar para cualquier vector $p := (p(1), p(2), \ldots, p(z))$ en $\mathbb{R}^z$, con $z \in \mathbb{N}$, definimos la norma l_∞, de la forma usual, es decir,

$$\|p\|_\infty = \max\{|p(1)|, |p(2)|, \ldots, |p(z)|\}. \tag{2.20}$$

Hipótesis A. Supongamos que las siguientes condiciones se cumplen:

(a) El espacio de acciones A es un espacio métrico compacto con métrica d_A.

(b) La función mediante la cuál se da la dinámica del contexto g es una función de Lipschitz continua con constante denotada por L, i.e., para $c, c' \in \mathbb{R}^d, m, m' \in \mathcal{P}(\mathcal{S}), a, a' \in A$

$$\|g(c, m, a) - g(c', m', a')\|_2 \leq L \max \{\|m - m'\|_1, \|c - c'\|_2, d_A(a, a')\}.$$

Sin pérdida de generalidad supondremos que dicha constante es mayor a 1, es decir, $L > 1$.

(c) La función $a \to H_\rho^N(\mathtt{y}, a, w)$ es continua para toda $\mathtt{y} \in \mathcal{Y}_N$ y $w \in \mathbb{R}^N$.

(d) Las funciones de costo ℓ y ℓ_T son acotadas por una constante $R > 0$ y uniformemente de Lipschitz con constante L_l. Es decir,

$$|\ell(\mathtt{y}, a)| + |\ell_T(\mathtt{y})| \leq R \quad \forall a \in A, \ \mathtt{y} \in \mathcal{Y}_N,$$

$$\sup_{(a,a') \in A \times A} |\ell(\mathtt{y}, a) - \ell(\mathtt{y}', a')| + |\ell_T(\mathtt{y}) - \ell_T(\mathtt{y}')| \leq L_l \|\mathtt{y} - \mathtt{y}'\| \quad \forall\, \mathtt{y}, \mathtt{y}' \in \mathcal{Y}_N, a, a' \in A.$$

Dado un espacio métrico genérico Z, denotamos por $\mathbb{M}(Z)$ al conjunto de funciones medibles $u : Z \mapsto \mathbb{R}$. Dicho lo anterior, comenzaremos con una proposición que viene siendo una consecuencia del teorema de selectores medibles. Para una demostración, se puede consultar la Proposición D5 en Hernández-Lerma y Lasserre [8].

Proposición 2.5.1. *Bajo la Hipótesis A, para cualquier función medible $u \in \mathbb{M}(\mathcal{Y}^N)$, se tiene que el mapeo*

$$\mathtt{y} \mapsto \inf_{a \in A} \left[\ell(\mathtt{y}, a) + \int_{\mathcal{Y}_N} u(x)\, Q_\rho(dx|\mathtt{y}, a) \right],$$

es medible. Además existe un selector medible $f \in \mathbb{F}^N$ tal que para cada $\mathtt{y} \in \mathcal{Y}^N$

$$\inf_{a \in A} \left[\ell(\mathtt{y}, a) + \int_{\mathcal{Y}_N} u(x)\, Q_\rho(dx|\mathtt{y}, a) \right] = \ell(\mathtt{y}, f(\mathtt{y})) + \int_{\mathcal{Y}_N} u(x)\, Q_\rho(dx|\mathtt{y}, f(\mathtt{y})).$$

Ahora bien, para toda $t = T-1, T-2, \ldots, 0$ y $\mathtt{y} \in \mathcal{Y}_N$, definimos la siguiente relación recursiva:

$$u_t(\mathtt{y}) = \inf_{a \in A} \left[\ell(\mathtt{y}, a) + \int_{\mathcal{Y}_N} u_{t+1}(x)\, Q_\rho(dx|\mathtt{y}, a) \right], \tag{2.21}$$

con condición terminal

$$u_T(\mathtt{y}) = \ell_T(\mathtt{y}). \tag{2.22}$$

Definimos también por $f_t \in \mathbb{F}^N$, $0 \leq t \leq T-1$ al selector medible tal que (recuerde la Proposición 2.5.1)

$$u_t(\mathtt{y}) = \ell(x, f_t(\mathtt{y})) + \int_{\mathcal{Y}_N} u_{t+1}(x) Q_\rho(dx|\mathtt{y}, f_t(\mathtt{y})) \quad \text{para } \mathtt{y} \in \mathcal{Y}_N. \tag{2.23}$$

Ahora consideremos la función de costo total del tiempo t en adelante: para $t = 0, 1, \ldots, T-1$, $\pi^N \in \Pi^N$,

$$C_t\left(\mathtt{h}_t^N, \pi^N\right) := E_{\mathtt{y}}^{\pi^N} \left[\sum_{s=t}^{T-1} \ell\left(\mathtt{Y}^N(s), \mathtt{a}(s)\right) + \ell_T\left(\mathtt{Y}^N(T)\right) \mid \mathtt{h}_t^N \right], \tag{2.24}$$

así como la función terminal

$$C_T(\mathtt{h}_T^N, \pi^N) := \ell_T(\mathtt{Y}^N(T)) \quad \forall\, \pi^N \in \Pi^N. \tag{2.25}$$

Finalmente considérese los costos óptimos:

$$U_t(\mathtt{h}_t^N) := \inf_{\pi^N \in \Pi^N} C_t^N\left(\mathtt{h}_t^N, \pi^N\right), \quad 0 \leq t \leq T, \tag{2.26}$$

Note que de las ecuaciones (2.22), (2.25) y (2.26), sobre $t = T$, se tiene

$$U_T(\mathtt{h}_T^N) = C_T(\mathtt{h}_T^N, \pi^N) = \ell_T(\mathtt{Y}^N(T)) = u_T(\mathtt{Y}^N(T)). \tag{2.27}$$

Nuestro objetivo a continuación será demostrar que $U_t(\mathtt{h}_t^N) = u_t(Y^N(t))$, para toda $t = 0, 1 \cdots, T-1$. Para ello usaremos inducción matemática hacia atrás.

A saber dado que tenemos que la relación ya se cumple para $t = T$ debido a (2.27), vamos a suponer que la relación se cumple para $t+1$ y vamos a probarla para t. Dada cualquier política $\pi^N \in \Pi^N$, se tiene que

$$\begin{aligned}
C_t\left(\mathtt{h}_t^N, \pi^N\right) &= E_{\mathtt{y}}^{\pi^N}\left[\sum_{s=t}^{T-1} \ell\left(\mathtt{Y}^N(s), \mathtt{a}(s)\right) + \ell_T\left(\mathtt{Y}^N(T)\right) \mid \mathtt{h}_t^N\right] \\
&= E_{\mathtt{y}}^{\pi^N}\left[\ell\left(\mathtt{Y}^N(t), \mathtt{a}(t)\right) + E_{\mathtt{y}}^{\pi^N}\left[\sum_{s=t+1}^{T-1} \ell\left(\mathtt{Y}^N(s), \mathtt{a}(s)\right) + \ell_T\left(\mathtt{Y}^N(T)\right) \mid \mathtt{h}_{t+1}^N\right] \mid \mathtt{h}_t^N\right] \\
&= E_{\mathtt{y}}^{\pi^N}\left[\ell\left(\mathtt{Y}^N(t), \mathtt{a}(t)\right) + C_{t+1}\left(\mathtt{h}_{t+1}^N, \pi^N\right) \mid \mathtt{h}_t^N\right] \\
&\geq E_{\mathtt{y}}^{\pi^N}\left[\ell\left(\mathtt{Y}^N(t), \mathtt{a}(t)\right) + U_{t+1}\left(\mathtt{h}_{t+1}^N\right) \mid \mathtt{h}_t^N\right] \quad (\text{por } (2.26)) \\
&= E_{\mathtt{y}}^{\pi^N}\left[\ell\left(\mathtt{Y}^N(t), \mathtt{a}(t)\right) + u_{t+1}\left(\mathtt{Y}^N(t+1)\right) \mid \mathtt{h}_t^N\right] \quad (\text{hipótesis de inducción}) \\
&= \int_A [\ell\left(\mathtt{Y}^N(t), a\right) + \int_{\mathcal{Y}_N} u_{t+1}(x) Q_\rho(dx|\mathtt{Y}^N(t), a)]\pi_t^N(da|\mathtt{h}_t^N) \\
&\geq u_t(\mathtt{Y}^N(t)).
\end{aligned}$$

Por otro lado, si insertamos la política determinística $\mathtt{f}_N^* = \{f_0^N, f_1^N, \cdots, f_{T-1}^N\}$ extraída de (2.23) en la expresión de arriba, podemos ver que se obtienen los mismos resultados pero con igualdades. Así, de lo anterior podemos deducir lo siguiente:

$$C_t\left(\mathtt{h}_t^N, \pi^N\right) \geq U_t\left(\mathtt{h}_t^N\right) \geq u_t(\mathtt{Y}^N(t)) \quad \text{y} \quad U_t\left(\mathtt{h}_t^N\right) = C_t\left(\mathtt{h}_t^N, \mathtt{f}_N^*\right) = u_t(\mathtt{Y}^N(t)). \tag{2.28}$$

A continuación mostraremos los resultados principales de esta sección.

Teorema 2.5.2. *Bajo los supuestos dados en la Hipótesis A, tenemos;*

(a) *La función de valor $V_*^N(y)$ definida en (2.17) coincide con la función $u_0(\mathtt{y})$ derivada de la iteración en (2.21). Además, dicha función es medible.*

(b) *La política $\mathtt{f}_N^*$ obtenida a través de (2.23), es óptima en el sentido de que ésta minimiza el lado derecho de (2.17).*

Demostración: La medibilidad de la función de valor $V_*^N(y)$ afirmada en (a), se sigue de usar iteradamente (T-veces) la Proposición 2.5.1 . El resto de la demostración se sigue de (2.28) cuando $t = 0$. $\square$

Observación 2.5.3. *Note que insertando la relación* (2.12) *a la EPD* (2.21) *se obtiene la siguiente ecuación, la cual es más explícita: para toda* $\mathtt{y} \in \mathcal{Y}_N$

$$\begin{aligned} u_t(\mathtt{y}) &= \inf_{a \in A} \left[\ell(\mathtt{y}, a) + \int_{\mathbb{R}^N} u_{t+1} \left(H_\rho^N(\mathtt{y}, a, w) \right) \theta(dw) \right], \\ u_T(\mathtt{y}) &= \ell_T(\mathtt{y}). \end{aligned} \tag{2.29}$$

Esto último implica que al ser N demasiado grande, la solubilidad de la ecuación de arriba puede resultar muy complicada debido a la dimensionalidad de $\mathbb{R}^N$, la cual podrá ser demasiado grande, ya que, justamente depende del número de partículas en interacción. En este sentido, se pueden hacer aproximaciones de este modelo que conllevará al estudio de un modelo límite el cual es denominado como un modelo de control de campo medio*.*

3

Modelo de control de campo medio

Como se comentó en el capítulo anterior, el hecho de demostrar la existencia de una política de control óptima, no es del todo suficiente para resolver nuestro problema ya que, desde el punto de vista aplicado, resulta ser que cuando el número de partículas N es muy grande, el problema se vuelve demasiado complejo.

Por ejemplo, para analizar las ecuaciones de optimalidad en (2.29), nos enfrentamos con una integral múltiple de dimensión N la cual podría resultar extremadamente díficil de calcular. Para superar este obstáculo recurriremos a la teoría de campo medio.

A grandes rasgos, primeramente introduciremos un modelo de control $\mathcal{M}$ que representará un modelo límite de $\mathcal{M}_N$ cuando $N \to \infty$; a este modelo lo llamaremos *modelo de control de campo medio*.

Una vez definido dicho modelo de control de campo medio, encontraremos condiciones de optimalidad del mismo. Más adelante, en capítulos subsecuentes, probaremos justo la anterior aseveración, es decir, mostraremos que $\mathcal{M}_N \to \mathcal{M}$ cuando $N \to \infty$ en un contexto adecuado. También buscaremos métodos de control para la construcción de las políticas óptimas con el objetivo de medir la diferencia entre el modelo original $\mathcal{M}_N$ y el modelo $\mathcal{M}$.

3.1 Descripción del modelo

Consideremos un sistema de control determinístico $\{(\mathtt{M}(t), \mathtt{C}(t))\}$ con valores en $\mathcal{Y} = \mathcal{P}(\mathcal{S}) \times \mathbb{R}^d$ con la siguiente dinámica:

$$\begin{aligned} \mathtt{M}(t+1) &= G_\rho(\mathtt{M}(t), \mathtt{C}(t), \mathtt{a}(t)), \\ \mathtt{C}(t+1) &= g(\mathtt{C}(t), \mathtt{M}(t+1), \mathtt{a}(t)). \end{aligned} \tag{3.1}$$

Con la condición inicial $(\mathtt{M}(0), \mathtt{C}(0)) = (\mathtt{m}_0, \mathtt{c}_0)$. Cabe señalar que $\mathtt{a}(t)$ es la acción tomada en el tiempo t por el agente controlador. Adicionalmente, la función g, la cual define la evolución del contexto del ambiente del sistema, es la función definida en la ecuación (2.5) y $G_\rho : \mathcal{P}(\mathcal{S}) \times \mathbb{R}^d \times A \to \mathcal{P}(\mathcal{S})$ es una función Lipschitz continua conocida y con constante de Lipschitz L_G, es decir, que para $c, c' \in \mathbb{R}^d, m, m' \in \mathcal{P}(\mathcal{S}), a, a' \in A$

$$\left\|G_\rho(m,c,a) - G_\rho(m',c',a')\right\|_1 \leq L_G \max\left\{\|m-m'\|_1, \|c-c'\|_2, d_A(a,a')\right\}. \tag{3.2}$$

Al ser un proceso determinístico, este mismo estará determinado por la sucesión de acciones (controles) $\{\mathtt{a}(t)\}$ así como por la condición inicial $(\mathtt{M}(0), \mathtt{C}(0)) \in \mathcal{Y}$.

Eventualmente supondremos que el proceso $\mathtt{Y}(t) := (\mathtt{M}(t), \mathtt{C}(t))$ representa el límite del N-MCM (2.11).

Ahora bien, al igual que en el N-MCM, definimos a la función $H_\rho : \mathcal{Y} \times A \to \mathcal{Y}$, conformada como sigue:

$$H_\rho(y,a) := \left(G_\rho(y,a), g\left(\mathtt{c}, G_\rho(y,a), a\right)\right) \quad \forall\, y = (m,c) \in \mathcal{Y},\ a \in A. \tag{3.3}$$

Esta función forma parte de la dinámica del proceso $\{(\mathtt{M}(t), \mathtt{C}(t))\}$, de la siguiente manera:

$$\begin{aligned} \mathtt{Y}(t+1) &= \left(G_\rho(\mathtt{Y}(t), \mathtt{a}(t)), g(\mathtt{Y}(t), \mathtt{a}(t))\right) \\ &= \left(G_\rho(\mathtt{Y}(t), \mathtt{a}(t)), g\left(\mathtt{C}(t), G_\rho(\mathtt{Y}(t), \mathtt{a}(t)), \mathtt{a}(t)\right)\right) \\ &= H_\rho(\mathtt{Y}(t), \mathtt{a}(t)). \end{aligned} \tag{3.4}$$

Bajo ciertas manipulaciones algebraicas podemos demostrar fácilmente que dicha función es de Lipschitz, por lo que, establecemos el siguiente resultado, el cual se demostrará inmediatamente:

Proposición 3.1.1. *Bajo la hipótesis en A, la función $H_\rho : \mathcal{Y} \times A \to \mathcal{Y}$ es de Lipschitz con constante $L_{H_\rho} = \max\{L_G, LL_G\}$.*

Demostración: Sean $(y,a), (y',a') \in \mathtt{Y} \times A$, donde $y = (m,c)$ y $y' = (m',c')$. Ahora bien, por el inciso (b) de la hipótesis A y de la ecuación (3.2) se obtiene lo siguiente:

$$\begin{aligned} \left\|G_\rho(m,c,a) - G_\rho(m',c',a')\right\|_1 &\leq L_G \max\left\{\|m-m'\|_1, \|c-c'\|_2, d_A(a,a')\right\}; \\ \left\|g(c,m,a) - g(c',m',a')\right\|_2 &\leq L \max\left\{\|m-m'\|_1, \|c-c'\|_2, d_A(a,a')\right\}. \end{aligned}$$

Ahora bien, de la ecuación (3.3) tenemos:

$$\begin{aligned} &\left\|H_\rho(y,a) - H_\rho(y',a')\right\|_\infty \\ &= \left\|\left(G_\rho(m,c,a), g\left(c, G_\rho(m,c,a), a\right)\right) - \left(G_\rho(m',c',a'), g\left(c', G_\rho(m',c',a'), a'\right)\right)\right\|_\infty \\ &= \max\left\{\left\|G_\rho(m,c,a) - G_\rho(m',c',a')\right\|_1, \left\|g\left(c, G_\rho(m,c,a), a\right) - g\left(c', G_\rho(m',c',a'), a'\right)\right\|_2\right\} \\ &\leq \max\{L_G, LL_G\} \max\left\{\|m-m'\|_1, \|c-c'\|_2, d_A(a,a')\right\} \\ &= \max\{L_G, LL_G\} \max\left\{\|y-y'\|, d_A(a,a')\right\}. \end{aligned}$$

Observemos que hemos hecho uso de la desigualdad que resulta de aplicar lo siguiente:

$$\begin{aligned} &\left\|g\left(c, G_\rho(m,c,a), a\right) - g\left(c, G_\rho(m',c',a'), a'\right)\right\|_2, \\ &\leq L \max\left\{\left\|G_\rho(m,c,a) - G_\rho(m',c',a')\right\|_1, \|c-c'\|_2, d_A(a,a')\right\}. \end{aligned}$$

Resumiendo, hemos obtenido lo siguiente:

$$\left\|H_\rho\left(y,a\right)-H_\rho\left(y',a'\right)\right\|_\infty \leq L_{H_\rho}\, \text{máx}\left\{\left\|y-y'\right\|, d_A\left(a,a'\right)\right\}. \tag{3.5}$$

Donde se ha definido:

$$L_{H_\rho}=\text{máx}\left\{L_G, LL_G\right\}.$$

$\square$

Dicho todo esto, y si consideramos ahora las funciones de costos establecidos en las definiciones (2.9) y (2.10) dentro del N-MCM (2.11), podemos definir el *modelo de control de campo medio* conforme a lo siguiente:

$$\mathcal{M}:=\left(\mathcal{Y},A,H_\rho,\ell,\ell_T\right). \tag{3.6}$$

Este modelo tiene una interpretación similar a la establecida en el N-MCM.

3.2 Planteamiento del problema de control

Ahora analizaremos los resultados ya conocidos relativos a la optimalidad del sistema que evoluciona conforme a las ecuaciones en (3.1), es decir, el modelo de control de campo medio (3.6).

Comenzaremos por revisar los conceptos de optimalidad para el modelo de campo medio (3.6). Su solución, la cual tiene que ver con la existencia de políticas óptimas y la caracterización de valores óptimos, se estudiará en la siguiente sección.

De forma análoga, que en el N modelo de control markoviano, definimos el espacio de historias al tiempo t como $\mathcal{H}_t:=\left(\mathcal{Y}\times A\right)^t\times\mathcal{Y}$ para $t\geq 1$ y $\mathcal{H}_0:=\mathcal{Y}$.

En contraste con la generalidad de políticas de la Definición 2.2.1, en este modelo solo consideraremos políticas de control *determinísticas* $\{f_t\}$ tales que $f_t\in\mathbb{F}$ para toda $t\in\mathbb{N}_0$ —ver la definición de $\mathbb{F}$ en el párrafo siguiente a la Definición 2.2.1(b); esto debido a dos razones: (1) por la naturaleza determinística del modelo de campo medio y (2) porque la políticas óptimas para el N-MCM se alcanzaron en las de tipo determinístico, es decir este último conjunto es un conjunto *suficiente* para hacer el análisis de optimalidad. Denotaremos por Π a este conjunto de políticas determinísticas $\pi=\{f_0,f_1,\cdots,f_{T-1}\}$.

Así que dada una política de control $\pi\in\Pi$, junto con la condición inicial $\mathtt{Y}\left(0\right)=\mathtt{y}\in\mathcal{Y}$ y con $T<\infty$, definiremos el costo total conforme a lo siguiente:

$$V\left(\pi,\mathtt{y}\right):=\sum_{t=0}^{T-1}\ell\left(\mathtt{Y}\left(t\right),\mathtt{a}\left(t\right)\right)+\ell_T\left(\mathtt{Y}\left(T\right)\right). \tag{3.7}$$

Adicionalmente, decimos que la política $\pi_*=\left\{f_0^*,f_1^*,f_2^*,\ldots,f_{T-1}^*\right\}\in\Pi$ es una política óptima para el modelo de control de campo medio (3.6) si y sólo sí satisface lo siguiente:

$$V\left(\pi_*,\mathtt{y}\right)=\inf_{\pi\in\Pi}V\left(\pi,\mathtt{y}\right):=V_*\left(\mathtt{y}\right)\qquad\forall\ \mathtt{y}\in\mathcal{Y}. \tag{3.8}$$

A la función V_* se le denomina la función de valor o costo óptimo del modelo de control de campo medio $\mathcal{M}$ y a π_* la política óptima. Cabe señalar que, de forma similar, podemos considerar el caso descontado al incluir únicamente un factor de descuento $0 < \alpha < 1$ y obtener así la siguiente función objetivo:

$$V(\pi, \mathtt{y}) := \sum_{t=1}^{T-1} \alpha^t \ell(\mathtt{Y}(t), \mathtt{a}(t)) + \alpha^T \ell_T(\mathtt{Y}). \tag{3.9}$$

3.3 Optimalidad en el Modelo de Campo Medio

Para obtener la optimalidad en este modelo, procederemos como en la sección anterior y usaremos la teoría de programación dinámica para generar los óptimos deseados. A saber, para toda $t = T-1, T-2, \ldots, 0$ y $\mathtt{y} \in \mathcal{Y}$, definimos la siguiente relación recursiva:

$$u_t(\mathtt{y}) = \inf_{a \in A} \left[\ell(\mathtt{y}, a) + u_{t+1}(H_\rho(\mathtt{y}, a))\right], \tag{3.10}$$

con condición terminal

$$u_T(\mathtt{y}) = \ell_T(\mathtt{y}). \tag{3.11}$$

En virtud de la Proposición 3.1.1 y de la Hipótesis A(d), los mapeos $a \mapsto H_\rho(\mathtt{y}, a)$ y $a \mapsto \ell(\mathtt{y}, a)$ son continuos y de esta manera se puede usar la Proposición 2.5.1 con sus respectivos cambios para garantizar que:

$$\mathtt{y} \mapsto \inf_{a \in A} \left[\ell(\mathtt{y}, a) + u(H_\rho(\mathtt{y}, a))\right],$$

es medible. Además existe un selector medible $f \in \mathbb{F}$ tal que, para cada $\mathtt{y} \in \mathcal{Y}$

$$\inf_{a \in A} \left[\ell(\mathtt{y}, a) + u(H_\rho(\mathtt{y}, a))\right] = \ell(\mathtt{y}, f(\mathtt{y})) + u(H_\rho(\mathtt{y}, f(\mathtt{y}))).$$

Esto implica que (3.10) y (3.11) estén bien definidos.

Para cada $t = 0, 1, 2, \ldots, T-1$, consideraremos al selector medible $f_t \in \mathbb{F}$ que minimiza el lador derecho de (3.10); es decir,

$$u_t(\mathtt{y}) = \ell(\mathtt{y}, f_t(\mathtt{y})) + u_{t+1}(H_\rho(\mathtt{y}, f_t(\mathtt{y}))). \tag{3.12}$$

Ahora tomemos la función de costo (3.7) del tiempo t en adelante para $t = 0, 1, \ldots, T-1$, $\pi \in \Pi$

$$C_t(\mathtt{h}_t, \pi) := \sum_{s=t}^{T-1} \ell(\mathtt{Y}(s), \mathtt{a}(s)) + \ell_T(\mathtt{Y}(T)), \tag{3.13}$$

donde hemos denotado por $\mathtt{h}_t := [\mathtt{Y}(0), \mathtt{a}(0), \cdots, \mathtt{Y}(t-1), \mathtt{a}(t-1), \mathtt{Y}(t)]$ a la historia generada desde el tiempo 0 hasta el tiempo t. Nuevamente se menciona que estas historias se diferencian unas de otras, únicamente por las diferentes configuraciones de las acciones, pues ya no existe aleatoriedad de por medio.

De igual manera, denotaremos la función terminal por

$$C_T(\mathtt{h}_T, \pi) := \ell_T(\mathtt{Y}(T)) \quad \forall\, \pi \in \Pi. \tag{3.14}$$

Finalmente considérese los costos óptimos

$$U_t(\mathtt{h}_t) := \inf_{\pi\in\Pi} C_t\left(\mathtt{h}_t, \pi\right), \quad 0 \leq t \leq T. \tag{3.15}$$

Note que de las ecuaciones (3.11), (3.14) y (3.15), sobre $t = T$, se tiene

$$U_T(\mathtt{h}_T) = C_T(\mathtt{h}_T, \pi) = \ell_T(\mathtt{Y}(T)) = u_T(\mathtt{Y}(T)). \tag{3.16}$$

Ahora bien, al igual que lo hicimos para el N-MCM, demostraremos que $U_t(\mathtt{h}_t) = u_t(\mathtt{Y}(t))$, para toda $t = 0, 1 \cdots, T-1$. Para ello usaremos inducción matemática hacia atrás.

A saber, dado que tenemos que la relación ya se cumple para $t = T$ debido a (3.16), vamos a suponer que la relación se cumple para $t+1$ y vamos a probarla para t. Asque dada cualquier política $\pi \in \Pi$, se tiene que

$$\begin{aligned}
C_t\left(\mathtt{h}_t, \pi\right) &= \sum_{s=t}^{T-1} \ell\left(\mathtt{Y}(s), \mathtt{a}(s)\right) + \ell_T\left(\mathtt{Y}(T)\right) \\
&= \ell\left(\mathtt{Y}(t), \mathtt{a}(t)\right) + \sum_{s=t+1}^{T-1} \ell\left(\mathtt{Y}(s), \mathtt{a}(s)\right) + \ell_T\left(\mathtt{Y}(T)\right) \\
&= \ell\left(\mathtt{Y}(t), \mathtt{a}(t)\right) + C_{t+1}\left(\mathtt{h}_{t+1}, \pi\right) \\
&\geq \ell\left(\mathtt{Y}(t), \mathtt{a}(t)\right) + U_{t+1}\left(\mathtt{h}_{t+1}\right) \quad \text{(por (3.15))} \\
&= \ell\left(\mathtt{Y}(t), \mathtt{a}(t)\right) + u_{t+1}\left(\mathtt{Y}(t+1)\right) \quad \text{(hipótesis de inducción)} \\
&= \ell\left(\mathtt{Y}(t), \mathtt{a}(t)\right) + u_{t+1}\left(H_\rho\left(\mathtt{Y}(t), \mathtt{a}(t)\right)\right) \quad \text{(por (3.4))} \\
&\geq u_t(\mathtt{Y}(t)).
\end{aligned} \tag{3.17}$$

Aplicando la política $\mathtt{f}^* = \{f_0, f_1, \cdots, f_{T-1}\}$ extraída de (3.12) en (3.17), podemos ver que se obtienen puras igualdades en el procedimiento de arriba. Así, de lo anterior podemos deducir lo siguiente:

$$C_t\left(\mathtt{h}_t, \pi\right) \geq U_t\left(\mathtt{h}_t\right) \geq u_t(\mathtt{Y}(t)) \quad \text{y} \quad C_t\left(\mathtt{h}_t, \mathtt{f}^*\right) = U_t\left(\mathtt{h}_t\right) = u_t(\mathtt{Y}(t)). \tag{3.18}$$

Y finalmente, al igual que en la sección previa, resumiremos los resultados, los cuales son análogos al N-modelo de control markoviano, pero ahora aplicados al modelo de control de campo medio:

Teorema 3.3.1. *Bajo los supuestos dados en la Hipótesis A, tenemos*

(a) *La función de valor $V_*(y)$ definida en (3.8) coincide con la función $u_0(\mathtt{y})$ derivada de la iteración en (3.10). Además, dicha función es medible.*

(b) *La política $\mathtt{f}^*$ obtenida a través de (3.12), es óptima en el sentido de que ésta minimiza el lado derecho de (3.8).*

Como podemos observar, tanto en el modelo de N-MCM como en el de campo medio, podemos garantizar la existencia de una política, sin embargo en el segundo mencionado hemos logrado eliminar el problema de dimensionalidad cuando $N \sim \infty$, lo cual es muy útil en las aplicaciones. En la siguiente sección mostraremos que el modelo de campo medio se puede ver como modelo límite del N-MCM y como consecuencia de esto, los resultados de optimalidad obtenidos en esta sección servirán para aproximar óptimos para el N-MCM.

4

Resultados de convergencia

Recordemos que en los capítulos anteriores se demostró la optimalidad del N-MCM $\mathcal{M}_N$ y la optimalidad asociada al modelo de control de campo medio $\mathcal{M}$. En este capítulo nos enfocaremos en estudiar la eficacia de la política óptima del modelo de control de campo medio $\mathcal{M}$ con respecto al N - modelo de control markoviano $\mathcal{M}_N$.

A fin de comprobar la eficacia del modelo de control de campo medio, mediremos la aproximación de la función de valor óptimo del modelo $\mathcal{M}_N$ con respecto a la del modelo de control de campo medio; es decir, la diferencia de las funciones V_*^N y V_*.

Para demostrar este hecho, impondremos hipótesis con respecto a la convergencia de las trayectorias de ambos procesos, es decir, las trayectorias del N-MCM $\left\{\mathtt{Y}^N(t)\right\}$ con respecto a las trayectorias del modelo de control de campo medio $\{\mathtt{Y}(t)\}$.

Cabe señalar que las políticas del N-MCM $\mathcal{M}_N$ recaen en Π^N y más aún, las óptimas son de tipo determinístico. Mientras que las políticas del modelo de campo medio $\mathcal{M}$ recaen en Π. De esta manera basta considerar para nuestro análisis de convergencia, políticas sobre el espacio Π.

Para tal fin, recordaremos algunos conceptos previos y añadiremos cierta notación que será necesaria a fin de obtener el resultado.

Primero notemos que el medir la diferencia de las dos trayectorias $\left\{\mathtt{Y}^N(t)\right\}$ y $\{\mathtt{Y}(t)\}$, será mediante la siguiente norma

$$\left\|\mathtt{Y}^N(t) - \mathtt{Y}(t)\right\| = \max\left\{\left\|\mathtt{M}^N(t) - \mathtt{M}(t)\right\|_1, \left\|\mathtt{C}^N(t) - \mathtt{C}(t)\right\|_2\right\},$$

la cual fue definida en capítulos anteriores. Adicional a esto, para cada política fija $\pi = \{f_t\} \in \Pi$, denotamos por:

$$\mathtt{a}^{\pi^N}(t) := f_t\left(\mathtt{Y}^N(t)\right), \quad \mathtt{a}^{\pi}(t) := f_t(\mathtt{Y}(t)), \tag{4.1}$$

a las acciones al tiempo t correspondientes a la aplicación de la política π bajo el proceso $\left\{\mathtt{Y}^N(t)\right\}$ y $\{\mathtt{Y}(t)\}$ respectivamente.

Finalmente, a fin de no incorporar notación demasiado grande, definimos los siguiente elementos para cada $T \in \mathbb{N}$

$$\begin{aligned} Y_T &:= \sup_{0\leq t\leq T} \left\| \mathtt{Y}^N(t) - \mathtt{Y}(t) \right\|; \\ \mathtt{K}(1) := L, \quad \text{y} \quad \mathtt{K}(T) &:= L^{T-1} \max\{L, diam(A)\} \quad \forall\, T \geq 2. \end{aligned} \tag{4.2}$$

Donde L es la constante de Lipschitz del inciso (b) de la Hipótesis A y además

$$diam(A) := \sup_{(\mathtt{a},\mathtt{a}')} d(\mathtt{a},\mathtt{a}').$$

Dicho esto, consideremos la siguiente hipótesis que daremos por válida en toda la sección. Cabe señalar que posteriormente demostraremos el inciso (b) de la misma dentro de las aplicaciones del siguiente capítulo.

Hipótesis B. Supongamos que las siguientes condiciones se cumplen:

(a) Para todo $N \in \mathbb{N}$ consideramos $\left(\mathtt{M}^N(0), \mathtt{C}^N(0)\right) = (\mathtt{M}(0), \mathtt{C}(0)) = (\mathtt{m}_0, \mathtt{c}_0) = \mathtt{y} \in \mathcal{Y}_N$.

(b) Para cada $\mathtt{y} \in \mathcal{Y}_N$, $T \in \mathbb{N}$ y $\varepsilon > 0$, existen constantes positivas K y λ tales que:

$$\sup_{\pi\in\Pi} \mathbb{P}^{\pi}_{\mathtt{y}} \left\{ \sup_{0\leq t\leq T} \left\| \mathtt{Y}^N(t) - \mathtt{Y}(t) \right\| \geq \gamma_T(\varepsilon) \right\} \leq KTe^{-\lambda N \varepsilon^2}, \tag{4.3}$$

donde $\gamma_T(\varepsilon) \to 0$ cuando $\varepsilon \to 0$.

Ahora bien, previo a establecer los resultados de convergencia principales de este capítulo (ver Teorema 4.0.3 más abajo), mostraremos dos proposiciones clave que nos permitirán concluir de forma más rápida y simple.

Proposición 4.0.1. *Para cada $\pi \in \Pi$ y $T \in \mathbb{N}$:*

(a)

$$Y_T := \sup_{0\leq t\leq T} \left\| \mathtt{Y}^N(t) - \mathtt{Y}(t) \right\| \leq \mathtt{K}(T). \tag{4.4}$$

(b) *Para cada $\mathtt{y} \in \mathcal{Y}_N$ y $T \in \mathbb{N}$ se tiene:*

$$\sup_{\pi\in\Pi} \mathbb{E}^{\pi}_{\mathtt{y}} \left[\sup_{0\leq t\leq T} \left\| \mathtt{Y}^N(t) - \mathtt{Y}(t) \right\| \right] \leq KTe^{-\lambda N \varepsilon^2} (1 + \mathtt{K}(T)) + \gamma_T(\varepsilon). \tag{4.5}$$

Demostración: Demostraremos primero el iniciso (a), para lo cual, notemos que bajo los supuestos de la Hipótesis B se sigue lo siguiente:

$$\mathtt{a}^{\pi^N}(0) = \mathtt{a}^{\pi}(0) = \mathtt{a}(0);$$

$$\left\| \mathtt{Y}^N(0) - \mathtt{Y}(0) \right\| = \|\mathtt{y} - \mathtt{y}\| = 0.$$

Hecho lo anterior, demostraremos este iniciso por medio del método de inducción matemática, así que, antes de realizar el paso inductivo, lo probaremos en particular para los casos $t = 1$ y $t = 2$.

Asimismo, observemos que si acotamos las diferencias de los procesos de proporciones del sistema, así como del contexto podremos acotar el de la pareja conformada por estos mismos.

Y dado que $\mathtt{M}^N(t) \in \mathcal{P}_N(\mathcal{S})$ y $\mathtt{M}(t) \in \mathcal{P}(\mathcal{S})$ son medidas de probabilidad, entonces se obtiene lo siguiente:

$$\forall t \geq 0 \quad \left\| \mathtt{M}^N(t) - \mathtt{M}(t) \right\|_1 \leq 1 \leq L. \tag{4.6}$$

Lo anterior debido al inciso (b) de la Hipótesis A, por lo que, la demostración del iniciso recae en los contextos de los procesos, observemos que en particular para $t = 1$ se obtiene:

$$\begin{aligned} \left\| \mathtt{C}^N(1) - \mathtt{C}(1) \right\|_2 &= \left\| g\left(\mathtt{C}^N(0), \mathtt{M}^N(1), \mathtt{a}^{\pi^N}(0) \right) - g\left(\mathtt{C}(0), \mathtt{M}(1), \mathtt{a}^{\pi}(0) \right) \right\|_2 \\ &= \left\| g\left(\mathtt{c}_0, \mathtt{M}^N(1), \mathtt{a}(0) \right) - g\left(\mathtt{c}_0, \mathtt{M}(1), \mathtt{a}(0) \right) \right\|_2 \\ &\leq L \max\left\{ \left\| \mathtt{M}^N(1) - \mathtt{M}(1) \right\|_1, \left\| \mathtt{c}_0 - \mathtt{c}_0 \right\|_2, d_A\left(\mathtt{a}(0), \mathtt{a}(0) \right) \right\} \\ &= L \left\| \mathtt{M}^N(1) - \mathtt{M}(1) \right\|_1. \end{aligned} \tag{4.7}$$

Observemos entonces que las ecuaciones (4.7) y (4.6) se resumen en la siguiente:

$$\left\| \mathtt{Y}^N(1) - \mathtt{Y}(1) \right\| = \max\left\{ \left\| \mathtt{M}^N(1) - \mathtt{M}(1) \right\|_1, \left\| \mathtt{C}^N(1) - \mathtt{C}(1) \right\|_2 \right\} \leq L.$$

Ahora para $t = 2$, se sigue lo siguiente para el contexto

$$\begin{aligned} \left\| \mathtt{C}^N(2) - \mathtt{C}(2) \right\|_2 &= \left\| g\left(\mathtt{C}^N(1), \mathtt{M}^N(2), \mathtt{a}^{\pi^N}(1) \right) - g\left(\mathtt{C}(1), \mathtt{M}(2), \mathtt{a}^{\pi}(1) \right) \right\|_2 \\ &\leq L \max\left\{ \left\| \mathtt{M}^N(2) - \mathtt{M}(2) \right\|_1, \left\| \mathtt{C}^N(1) - \mathtt{C}(1) \right\|_2, d_A\left(\mathtt{a}^{\pi^N}(1), \mathtt{a}^{\pi}(1) \right) \right\} \\ &\leq L \max\left\{ 1, L, diam(A) \right\} \leq L \max\left\{ L, diam(A) \right\} = L\mathtt{K}(1) = \mathtt{K}(2). \end{aligned} \tag{4.8}$$

Y de forma análoga se sigue ahora por las ecuaciones (4.8) y (4.6) que:

$$\left\| \mathtt{Y}^N(2) - \mathtt{Y}(2) \right\| = \max\left\{ \left\| \mathtt{M}^N(2) - \mathtt{M}(2) \right\|_1, \left\| \mathtt{C}^N(2) - \mathtt{C}(2) \right\|_2 \right\} \leq \mathtt{K}(2).$$

Por lo que, en general supongamos, para algún $t \in \mathbb{N}$ lo siguiente:

$$\left\| \mathtt{Y}^N(t) - \mathtt{Y}(t) \right\| = \max\left\{ \left\| \mathtt{M}^N(t) - \mathtt{M}(t) \right\|_1, \left\| \mathtt{C}^N(t) - \mathtt{C}(t) \right\|_2 \right\} \leq \mathtt{K}(t). \tag{4.9}$$

A partir de lo anterior, consideremos ahora el caso para $t+1$, a fin de establecer el resultado y debido a la ecuación (4.6) nos centraremos en el contexto, así que

$$\begin{aligned} &\left\| \mathtt{C}^N(t+1) - \mathtt{C}(t+1) \right\|_2 = \left\| g\left(\mathtt{C}^N(t), \mathtt{M}^N(t+1), \mathtt{a}^{\pi^N}(t) \right) - g\left(\mathtt{C}(t), \mathtt{M}(t+1), \mathtt{a}^{\pi}(t) \right) \right\|_2 \\ &\leq L \max\left\{ \left\| \mathtt{M}^N(t+1) - \mathtt{M}(t+1) \right\|_1, \left\| \mathtt{C}^N(t) - \mathtt{C}(t) \right\|_2, d_A\left(\mathtt{a}^{\pi^N}(t), \mathtt{a}^{\pi}(t) \right) \right\} \\ &\leq L \max\left\{ 1, \mathtt{K}(t), diam(A) \right\} = L \max\left\{ \mathtt{K}(t), diam(A) \right\} \\ &= L \max\left\{ L^{t-1} \max\left\{ L, diam(A) \right\}, diam(A) \right\} = L^t \max\left\{ L, diam(A) \right\} = \mathtt{K}(t+1). \end{aligned} \tag{4.10}$$

Así que por las ecuaciones (4.10) y (4.6) se obtiene el resultado buscado.

Ahora demostraremos el inciso (b) de la proposición, para lo cual considere $\mathtt{y} \in \mathcal{Y}_N$, $\pi \in \Pi$, $T \in \mathbb{N}$ y $\varepsilon > 0$, entonces por las propiedades de la esperanza se sigue que

$$\begin{aligned}
\mathbb{E}_{\mathtt{y}}^{\pi}\left[Y_T\right] &= \mathbb{E}_{\mathtt{y}}^{\pi}\left[Y_T \mathbb{1}_{\{Y_T \geq \gamma_T(\varepsilon)\}} + Y_T \mathbb{1}_{\{Y_T < \gamma_T(\varepsilon)\}}\right] \\
&\leq \mathbb{E}_{\mathtt{y}}^{\pi}\left[Y_T \mathbb{1}_{\{Y_T \geq \gamma_T(\varepsilon)\}}\right] + \gamma_T(\varepsilon)\, \mathbb{P}_{\mathtt{y}}^{\pi}\left(Y_T < \gamma_T(\varepsilon)\right) \\
&\leq \mathbb{E}_{\mathtt{y}}^{\pi}\left[Y_T \mathbb{1}_{\{Y_T \geq \gamma_T(\varepsilon)\}}\right] + \gamma_T(\varepsilon) \\
&\leq \left(1 + \mathtt{K}(T)\right) \mathbb{P}_{\mathtt{y}}^{\pi}\left(Y_T \geq \gamma_T(\varepsilon)\right) + \gamma_T(\varepsilon) \\
&\leq \left(1 + \mathtt{K}(T)\right) \mathbb{P}_{\mathtt{y}}^{\pi}\left\{\sup_{0 \leq t \leq T} \left\|\mathtt{Y}^N(t) - \mathtt{Y}(t)\right\| \geq \gamma_T(\varepsilon)\right\} + \gamma_T(\varepsilon) \\
&\leq \left(1 + \mathtt{K}(T)\right) K T e^{-\lambda N \varepsilon^2} + \gamma_T(\varepsilon).
\end{aligned}$$

Ahora bien, notemos que la tercera desigualdad se debe al hecho que $Y_T \leq \mathtt{K}(T)$ obteniendo así

$$\frac{Y_T}{1 + \mathtt{K}(T)} \leq \frac{Y_T}{1 + Y_T} \leq 1,$$

por lo que, debido a esto se obtiene la siguiente desigualdad

$$\frac{1}{1 + \mathtt{K}(T)} Y_T \mathbb{1}_{\{Y_T \geq \gamma_T(\varepsilon)\}} \leq \mathbb{1}_{\{Y_T \geq \gamma_T(\varepsilon)\}}.$$

Adicionalmente, la quinta desigualdad se sigue de la Hipótesis B, con lo que hemos obtenido hasta el momento.

$$\mathbb{E}_{\mathtt{y}}^{\pi}\left[Y_T\right] \leq \left(1 + \mathtt{K}(T)\right) K T e^{-\lambda N \varepsilon^2} + \gamma_T(\varepsilon).$$

Y al tomar el supremo sobre $\pi \in \Pi$ se sigue el resultado deseado. □

Notemos que a partir de la proposición anterior, se ha establecido una cota para la diferencia de las normas de las trayectorias del N-MCM con respecto al modelo de campo medio.

Así que, ahora nos centraremos en buscar una cota para la diferencia de las funciones de costo por etapa, así como de las funciones objetivo bajo cierta política. Para esto, establecemos la siguiente proposición.

Proposición 4.0.2. *Sea L_l la constante de la Hipótesis A(d). Entonces, para cada $\mathtt{y} \in \mathcal{Y}_N$, $\varepsilon > 0$, $T \in \mathbb{N}$ y $0 \leq t \leq T$, las siguientes relaciones se cumplen:*

(a)

$$\sup_{\pi \in \Pi} \mathbb{E}_{\mathtt{y}}^{\pi}\left[\left|\ell\left(\mathtt{Y}^N(t), \mathtt{a}^{\pi^N}(t)\right) - \ell\left(\mathtt{Y}(t), \mathtt{a}^{\pi}(t)\right)\right|\right] \leq L_l\left(K T e^{-\lambda N \varepsilon^2}\left(1 + \mathtt{K}(T)\right) + \gamma_T(\varepsilon)\right), \tag{4.11}$$

(b)

$$\mathbb{E}_{\mathtt{y}}^{\pi}\left[\sup_{\pi \in \Pi}\left|V^N\left(\pi, \mathtt{Y}^N(t)\right) - V\left(\pi, \mathtt{Y}(t)\right)\right|\right] \leq L_l(T+1)\left(K T e^{-\lambda N \varepsilon^2}\left(1 + \mathtt{K}(T)\right) + \gamma_T(\varepsilon)\right). \tag{4.12}$$

Demostración: Comenzaremos por demostrar el inciso (a), por lo que, fijemos una política $\pi \in \Pi$ y $T \in \mathbb{N}$. Entonces por la Hipótesis A(d) se sigue:

$$\begin{aligned}
&\mathbb{E}_{\mathtt{y}}^{\pi}\left[\left|\ell\left(\mathtt{Y}^N(t), \mathtt{a}^{\pi^N}(t)\right) - \ell\left(\mathtt{Y}(t), \mathtt{a}^{\pi}(t)\right)\right|\right] + \mathbb{E}_{\mathtt{y}}^{\pi}\left[\left|\ell_T\left(\mathtt{Y}^N(t)\right) - \ell_T\left(\mathtt{Y}(t)\right)\right|\right] \\
&\leq L_l \mathbb{E}_{\mathtt{y}}^{\pi}\left[\left\|\mathtt{Y}^N(t) - \mathtt{Y}(t)\right\|\right] \leq L_l\left(KTe^{-\lambda N\varepsilon^2}\left(1 + \mathtt{K}(T)\right) + \gamma_T(\varepsilon)\right).
\end{aligned}$$

Y al tomar el supremo sobre $\pi \in \Pi$ se obtiene el resultado deseado.

Ahora bien, continuaremos con demostrar el inciso (b). Así pues, escogemos $\pi \in \Pi$, de tal manera que

$$\begin{aligned}
&\left|V^N\left(\pi, \mathtt{Y}^N(t)\right) - V\left(\pi, \mathtt{Y}(t)\right)\right| \\
&= \left|E_{\mathtt{Y}^N(t)}^{\pi^N}\left[\sum_{k=t}^{T-1} \ell\left(\mathtt{Y}^N(k), \mathtt{a}(k)\right) + \ell_T\left(\mathtt{Y}^N(T)\right)\right] - \sum_{k=t}^{T-1} \ell\left(\mathtt{Y}(k), \mathtt{a}(k)\right) - \ell_T\left(\mathtt{Y}(T)\right)\right| \\
&= \left|E_{\mathtt{Y}^N(t)}^{\pi^N}\left[\sum_{k=t}^{T-1} \ell\left(\mathtt{Y}^N(k), \mathtt{a}(k)\right) + \ell_T\left(\mathtt{Y}^N(T)\right) - \sum_{k=t}^{T-1} \ell\left(\mathtt{Y}(k), \mathtt{a}(k)\right) - \ell_T\left(\mathtt{Y}(T)\right)\right]\right| \\
&= \left|E_{\mathtt{Y}^N(t)}^{\pi^N}\left[\sum_{k=t}^{T-1}\left(\ell\left(\mathtt{Y}^N(k), \mathtt{a}(k)\right) - \ell\left(\mathtt{Y}(k), \mathtt{a}(k)\right)\right) + \left(\ell_T\left(\mathtt{Y}^N(T)\right) - \ell_T\left(\mathtt{Y}(T)\right)\right)\right]\right| \\
&= \left|\sum_{k=t}^{T-1} E_{\mathtt{Y}^N(t)}^{\pi^N}\left[\ell\left(\mathtt{Y}^N(k), \mathtt{a}(k)\right) - \ell\left(\mathtt{Y}(k), \mathtt{a}(k)\right)\right] + E_{\mathtt{Y}^N(t)}^{\pi^N}\left[\ell_T\left(\mathtt{Y}^N(T)\right) - \ell_T\left(\mathtt{Y}(T)\right)\right]\right| \\
&\leq \sum_{k=t}^{T-1} E_{\mathtt{Y}^N(k)}^{\pi^N}\left[\left|\ell\left(\mathtt{Y}^N(k), \mathtt{a}(k)\right) - \ell\left(\mathtt{Y}(k), \mathtt{a}(k)\right)\right|\right] + E_{\mathtt{Y^N(t)}}^{\pi^N}\left[\left|\ell_T\left(\mathtt{Y}^N(T)\right) - \ell_T\left(\mathtt{Y}(T)\right)\right|\right] \\
&\leq \sum_{k=t}^{T-1} E_{\mathtt{Y}^N(k)}^{\pi^N}\left[L_l\left\|\mathtt{Y}^N(k) - \mathtt{Y}(k)\right\|\right] + E_{\mathtt{Y}^N(t)}^{\pi^N}\left[L_l\left\|\mathtt{Y}^N(T) - \mathtt{Y}(T)\right\|\right] \\
&\leq \sum_{k=t}^{T-1} L_l E_{\mathtt{Y}^N(k)}^{\pi^N}\left[\sup_{t\leq k\leq T}\left\|\mathtt{Y}^N(k) - \mathtt{Y}(k)\right\|\right] + L_l E_{\mathtt{Y}^N(t)}^{\pi^N}\left[\sup_{t\leq k\leq T}\left\|\mathtt{Y}^N(k) - \mathtt{Y}(k)\right\|\right] \\
&\leq \sum_{t=0}^{T-1} L_l\left(KTe^{-\lambda N\varepsilon^2}\left(1 + \mathtt{K}(T)\right) + \gamma_T(\varepsilon)\right) + L_l\left(KTe^{-\lambda N\varepsilon^2}\left(1 + \mathtt{K}(T)\right) + \gamma_T(\varepsilon)\right) \\
&= L_l(T+1)\left(KTe^{-\lambda N\varepsilon^2}\left(1 + \mathtt{K}(T)\right) + \gamma_T(\varepsilon)\right).
\end{aligned}$$

Así que entonces, hemos obtenido la siguiente desigualdad

$$\left|V^N\left(\pi, \mathtt{Y}^N(t)\right) - V\left(\pi, \mathtt{Y}(t)\right)\right| \leq L_l(T+1)\left(KTe^{-\lambda N\varepsilon^2}\left(1 + \mathtt{K}(T)\right) + \gamma_T(\varepsilon)\right). \tag{4.13}$$

Finalmente, tomando primero el supremo sobre $\pi \in \Pi$ y posteriormente la esperanza se obtiene el resultado. $\square$

Observemos que a partir de estos resultados se puede apreciar que la diferencia de las funciones de costo en cada etapa, con respecto a la evolución de ambos modelos se encuentra acotada.

Más aún, observemos que la diferencia de las funciones objetivo de ambos sistemas se encuentran acotadas, por lo que, ya contamos con los elementos necesarios para formalizar el resultado principal de esta sección que a continuación se enuncia.

Teorema 4.0.3. *Supongamos que se cumplen las Hipótesis A y B. Entonces, para cada $T \in \mathbb{N}$, $0 \leq t \leq T$ y $\mathtt{y} \in \mathcal{Y}_N$ se satisface lo siguiente:*

$$\mathbb{E}_{\mathtt{y}}^{\pi}\left[\left|V_*^N\left(\mathtt{Y}^N(t)\right) - V_*\left(\mathtt{Y}(t)\right)\right|\right] \leq L_l\left(T+1\right)\left(KTe^{-\lambda N\varepsilon^2}\left(1+\mathtt{K}\left(T\right)\right) + \gamma_T\left(\varepsilon\right)\right). \tag{4.14}$$

Demostración: Sea $\pi_*^N = \left\{f_0^{N*}, \cdots, f_{T-1}^{N*}\right\} \in \Pi^N$ una política óptima para el N-MCM $\mathcal{M}_N$ —la cual se puede escoger del tipo determinística (ver Teorema 2.5.2)— y para algún selector arbitrario $\overline{f} : \mathcal{Y} \to A$, el cual esta contenido en $\mathbb{F}$, definimos la política $\overline{\pi} = \left\{\overline{f}_0, \cdots, \overline{f}_{T-1}\right\} \in \Pi$, donde:

$$\overline{f}_t\left(\mathtt{y}\right) := f_t^{N*}\left(\mathtt{y}\right)\mathbb{1}_{\mathcal{Y}_N}\left(\mathtt{y}\right) + \overline{f}\left(\mathtt{y}\right)\mathbb{1}_{\mathcal{Y}_N^c}\left(\mathtt{y}\right).$$

Observemos que para cada $0 \leq t \leq T$ se tiene

$$V_*^N\left(\mathtt{Y}^N(t)\right) = V^N\left(\pi_*^N, \mathtt{Y}^N(t)\right) = V^N\left(\overline{\pi}, \mathtt{Y}^N(t)\right) \leq \sup_{\pi\in\Pi} V^N\left(\pi, \mathtt{Y}^N(t)\right).$$

Entonces,

$$V_*^N\left(\mathtt{Y}^N(t)\right) - V_*\left(\mathtt{Y}(t)\right) \leq \sup_{\pi\in\Pi} V^N\left(\pi, \mathtt{Y}^N(t)\right) - \inf_{\pi\in\Pi} V\left(\pi, \mathtt{Y}(t)\right),$$

y por propiedades del supremo e ínfimo se obtiene lo siguiente

$$\left|V_*^N\left(\mathtt{Y}^N(t)\right) - V_*\left(\mathtt{Y}(t)\right)\right| \leq \sup_{\pi\in\Pi}\left|V^N\left(\pi, \mathtt{Y}^N(t)\right) - V\left(\pi, \mathtt{Y}(t)\right)\right|.$$

Además sea $\varphi \in \Pi$ una política arbitaria y denotamos por $\mathtt{Y}_\varphi^N(t) = \mathtt{Y}^N(t) \in \mathcal{Y}_N$ y $\mathtt{Y}_\varphi(t) = \mathtt{Y}(t) \in \mathcal{Y}$. Luego entonces, para cada $\mathtt{y} \in \mathcal{Y}_N$ y $0 \leq t \leq T$, se sigue

$$\begin{aligned}\mathbb{E}_{\mathtt{y}}^{\varphi}\left[\left|V_*^N\left(\mathtt{Y}^N(t)\right) - V_*\left(\mathtt{Y}(t)\right)\right|\right] &\leq \mathbb{E}_{\mathtt{y}}^{\varphi}\left[\sup_{\pi\in\Pi}\left|V^N\left(\pi, \mathtt{Y}^N(t)\right) - V\left(\pi, \mathtt{Y}(t)\right)\right|\right] \\ &\leq L_l\left(T+1\right)\left(KTe^{-\lambda N\varepsilon^2}\left(1+\mathtt{K}\left(T\right)\right) + \gamma_T\left(\varepsilon\right)\right),\end{aligned}$$

con lo que queda demostrado el resultado. □

Ahora bien, observemos que con este teorema hemos podido establecer que la diferencias entre la función de valor óptima del N- modelo de control markoviano $\mathcal{M}_N$ y la función de valor del modelo de control de campo medio $\mathcal{M}$ óptima se pueden acotar.

Así que, a partir de este resultado podemos concluir que a fin de evitar el problema de dimensionalidad, generados al emplear y trabajar con el modelo $\mathcal{M}_N$, podemos trabajar con el modelo de control de campo medio $\mathcal{M}$.

Más aún, si determinamos una política de control óptima en el modelo de control de campo medio y determinamos el valor óptimo de la función objetivo, obtenido al evaluar dicha política, entonces esta misma aproxima el valor óptimo del N - modelo de control markoviano $\mathcal{M}_N$.

5

Aplicaciones

A fin de ilustrar los resultados obtenidos en este trabajo, presentaremos dos aplicaciones en las que intervienen la interacción de un número significativo de partículas u objetos; estos últimos tendrán una representación específica en cada aplicación. El análisis consistirá en mostrar que las hipótesis asumidas en la teoría se satisfacen en estas instancias y así, garantizar la existencia de políticas óptimas en el modelo de campo medio que servirán como aproximaciones del modelo original en cuestión.

5.1 Modelo de consumo-inversión

Consideraremos un sistema de consumo-inversión compuesto por N inversionistas (partículas), en el cual supondremos que las decisiones individuales de cada uno de estos inversionistas no tienen influencia en los precios del mercado. Estos agentes invierten su capital en diversos activos con diferentes tasas de retorno (rendimiento) y además consumen algún producto en particular.

Además supondremos la existencia de un agente controlador, el cual puede ser el gobierno del país o alguna otra entidad reguladora, que puede realizar dos diferentes actividades:

1. Subsidiar cierta porción de las inversiones, a fin de incentivar y motivar la inversión.
2. Decretar la implementación de un impuesto sobre las inversiones.

A fin de no complicar los cálculos, supondremos que sólo existen dos activos: uno sin riesgo a una tasa de retorno fija τ y otro con riesgo, con una tasa de retorno aleatoria $\{\xi_t\}$, la cual toma valores en un conjunto acotado $Z \subseteq \mathbb{R}$.

De acuerdo con lo expresado previamente, cada inversionista puede invertir en ambos activos (con y sin riesgo). De esta manera, denotaremos por $\varphi_1 : \mathbb{R}^d \to [0, 1]$ a la proporción del capital invertido en el activo con riesgo, misma que dependerá del contexto del ambiente; este ambiente puede ser, por ejemplo, incertidumbre de los inversionistas y/o los tipos de mercado en los que se esté invirtiendo y/o frecuencias de transacciones, etc. Por otro lado, la proporción $(1 - \varphi_1)$ será la fracción del capital invertido en activo libre de riesgo.

Además de invertir, supondremos que cada inversionista consume una cantidad del capital que dispone; a esta cantidad la denotaremos por la función $\varphi_2 : \mathbb{R}^d \to \mathbb{R}^+$, la cual será una función acotada que también dependerá del contexto del ambiente.

En este ejemplo, el espacio de estados consiste en los posibles grados de riqueza (en montos) que los N inversionistas pueden alcanzar. En específico, consideraremos estos grados como el conjunto finito $\mathcal{S} := \{0, 1, 2, \ldots, S\}$. Al considerar el conjunto anterior con componentes enteras, estaremos despreciando los centavos.

Sea $\mathtt{a}(t) \in A := \{0, \pm 1, \pm 2, \pm 3, \ldots, \pm \mathtt{a}^*\}$, para algún $\mathtt{a}^* \geq 0$, las acciones del agente controlador las cuales tienen la siguiente interpretación:

- Siempre que $\mathtt{a}(t) < 0$, implicará que el agente controlador impuso un impuesto de valor $-\mathtt{a}(t)$.
- Siempre que $\mathtt{a}(t) > 0$, implicará que el agente controlador impuso un subsidio de valor $\mathtt{a}(t)$.

Ahora denotaremos por $X_n^N(t) \in \mathcal{S} := \{0, 1, 2, \ldots, S\}$ el capital del inversionista n al tiempo t. Con los componentes anteriores, la evolución del sistema se encontrará dada por

$$X_n^N(t+1) = F\left(X_n^N(t), \mathtt{C}^N(t), \mathtt{a}(t), \xi_t\right). \tag{5.1}$$

Donde la función de transición de la dinámica se define como

$$F(i, \mathtt{c}, \mathtt{a}, z) := \mathtt{e}\left\{\left[\left(1 - \varphi_1(\mathtt{c})\right)(1+\tau) + \varphi_1(\mathtt{c})\, z\right]\left[i - \varphi_2(\mathtt{c}) + \mathtt{a}\right]\right\}. \tag{5.2}$$

Donde denotamos a $\mathtt{e}\{x\}$ a la parte entera (función suelo) de x.

En toda esta aplicación, vamos a suponer que las funciones φ_1 y φ_2 son funciones de Lipschitz con constantes L_{φ_1} y L_{φ_2} respectivamente.

Proposición 5.1.1. *La función F definida en (5.2) es Lipschitz en las variables de acción y del contexto del ambiente, con constante*

$$L_F := 1 + (1 + \tau + z^*)\left(iL_{\varphi_1} + \overline{L}_{\varphi_1} L_{\varphi_2} + \overline{L}_{\varphi_2} L_{\varphi_1} + \mathtt{a}^* L_{\varphi_1} + L_{\varphi_2}\right) + (1+\tau)\left(1 + L_{\varphi_2}\right),$$

donde $\overline{L}_{\varphi_1}$ y $\overline{L}_{\varphi_2}$ son las cotas de las funciones φ_1 y φ_2 respectivamente, además $\mathtt{a}^* = \max_{\mathtt{a} \in A} |a|$ y $z^* := \max_{z \in Z} |z|$.

Demostración: Primero señalaremos una desigualdad válida para la función parte entera y para cualesquiera números reales a y b, la cual, usaremos posteriormente

$$\begin{aligned} |\mathtt{e}\{a\} - \mathtt{e}\{b\}| &= |(a - \alpha) - (b - \beta)| = |(a - b) + (\beta - \alpha)| \leq |a - b| + |\beta - \alpha| \\ &\leq |a - b| + 1, \end{aligned}$$

donde α, β satisface la relación $a - e\{a\} = \alpha$ y $b - e\{b\} = \beta$.

Ahora nos centraremos en demostrar que la función F es de Lipschitz en las variables de acción y del contexto del ambiente. Así pues, sean $(\mathsf{a},\mathsf{c}),(\mathsf{a}',\mathsf{c}')\in A\times\mathbb{R}^d$, e $i\in\mathcal{S}$ y $z\in\mathbb{R}$ fijos, entonces se tiene:

$$
\begin{aligned}
&|F(i,\mathsf{c},\mathsf{a},z)-F(i,\mathsf{c}',\mathsf{a}',z)|=|\mathsf{e}\{[(1-\varphi_1(\mathsf{c}))(1+\tau)+\varphi_1(\mathsf{c})z][i-\varphi_2(\mathsf{c})+\mathsf{a}]\}\\
&-\mathsf{e}\{[(1-\varphi_1(\mathsf{c}'))(1+\tau)+\varphi_1(\mathsf{c}')z][i-\varphi_2(\mathsf{c}')+\mathsf{a}']\}|\\
&\leq 1+|[(1-\varphi_1(\mathsf{c}))(1+\tau)+\varphi_1(\mathsf{c})z][i-\varphi_2(\mathsf{c})+\mathsf{a}]\\
&-[(1-\varphi_1(\mathsf{c}'))(1+\tau)+\varphi_1(\mathsf{c}')z][i-\varphi_2(\mathsf{c}')+\mathsf{a}']|\\
&=1+|(1-\varphi_1(\mathsf{c}))(1+\tau)i+(1-\varphi_1(\mathsf{c}))(1+\tau)(-\varphi_2(\mathsf{c})+\mathsf{a})+\varphi_1(\mathsf{c})zi\\
&+\varphi_1(\mathsf{c})z(-\varphi_2(\mathsf{c})+\mathsf{a})-(1-\varphi_1(\mathsf{c}'))(1+\tau)i-(1-\varphi_1(\mathsf{c}'))(1+\tau)(-\varphi_2(\mathsf{c}')+\mathsf{a}')\\
&-\varphi_1(\mathsf{c}')zi-\varphi_1(\mathsf{c}')z(-\varphi_2(\mathsf{c}')+\mathsf{a}')|\\
&=1+|(1+\tau)i(1-\varphi_1(\mathsf{c})-1+\varphi_1(\mathsf{c}'))+zi(\varphi_1(\mathsf{c})-\varphi_1(\mathsf{c}'))\\
&+z[\varphi_1(\mathsf{c})(-\varphi_2(\mathsf{c})+\mathsf{a})-\varphi_1(\mathsf{c}')(-\varphi_2(\mathsf{c}')+\mathsf{a}')]\\
&+(1+\tau)[(1-\varphi_1(\mathsf{c}))(-\varphi_2(\mathsf{c})+\mathsf{a})-(1-\varphi_1(\mathsf{c}'))(-\varphi_2(\mathsf{c}')+\mathsf{a}')]|\\
&\leq 1+(1+\tau+z^*)iL_{\varphi_1}\|\mathsf{c}-\mathsf{c}'\|_2\\
&+z^*|-\varphi_1(\mathsf{c})\varphi_2(\mathsf{c})+\varphi_1(\mathsf{c})\mathsf{a}+\varphi_1(\mathsf{c}')\varphi_2(\mathsf{c}')-\varphi_1(\mathsf{c}')\mathsf{a}'|\\
&+(1+\tau)|(-\varphi_2(\mathsf{c})+\mathsf{a})-\varphi_1(\mathsf{c})(-\varphi_2(\mathsf{c})+\mathsf{a})-(-\varphi_2(\mathsf{c}')+\mathsf{a}')\\
&+\varphi_1(\mathsf{c}')(-\varphi_2(\mathsf{c}')+\mathsf{a}')|\\
&\leq 1+(1+\tau+z^*)iL_{\varphi_1}\|\mathsf{c}-\mathsf{c}'\|_2+(1+\tau)|-\varphi_2(\mathsf{c})+\mathsf{a}+\varphi_2(\mathsf{c}')-\mathsf{a}'|\\
&+z^*|\varphi_1(\mathsf{c})\varphi_2(\mathsf{c})+\varphi_1(\mathsf{c})\mathsf{a}+\varphi_1(\mathsf{c}')\varphi_2(\mathsf{c}')-\varphi_1(\mathsf{c}')\mathsf{a}'|\\
&+(1+\tau)|\varphi_1(\mathsf{c})\varphi_2(\mathsf{c})-\varphi_1(\mathsf{c})\mathsf{a}-\varphi_1(\mathsf{c}')\varphi_2(\mathsf{c}')+\varphi_1(\mathsf{c}')\mathsf{a}'|\\
&\leq 1+(1+\tau+z^*)iL_{\varphi_1}\|\mathsf{c}-\mathsf{c}'\|_2+(1+\tau)|\varphi_2(\mathsf{c}')-\varphi_2(\mathsf{c})|+(1+\tau)|\mathsf{a}-\mathsf{a}'|\\
&+(1+\tau+z^*)|\varphi_1(\mathsf{c})\varphi_2(\mathsf{c})-\varphi_1(\mathsf{c})\mathsf{a}-\varphi_1(\mathsf{c}')\varphi_2(\mathsf{c}')+\varphi_1(\mathsf{c}')\mathsf{a}'|\\
&\leq 1+(1+\tau+z^*)iL_{\varphi_1}\|\mathsf{c}-\mathsf{c}'\|_2+(1+\tau)L_{\varphi_2}\|\mathsf{c}-\mathsf{c}'\|_2+(1+\tau)d_A(\mathsf{a},\mathsf{a}')\\
&+(1+\tau+z^*)|\varphi_1(\mathsf{c})\varphi_2(\mathsf{c})-\varphi_1(\mathsf{c}')\varphi_2(\mathsf{c}')|+(1+\tau+z^*)|-\varphi_1(\mathsf{c})\mathsf{a}+\varphi_1(\mathsf{c}')\mathsf{a}'|\\
&\leq 1+(1+\tau+z^*)\left(iL_{\varphi_1}+\overline{L}_{\varphi_1}L_{\varphi_2}+\overline{L}_{\varphi_2}L_{\varphi_1}\right)\|\mathsf{c}-\mathsf{c}'\|_2+(1+\tau)L_{\varphi_2}\|\mathsf{c}-\mathsf{c}'\|_2\\
&+(1+\tau)d_A(\mathsf{a},\mathsf{a}')+(1+\tau+z^*)\mathsf{a}^*L_{\varphi_1}\|\mathsf{c}-\mathsf{c}'\|_2\\
&\leq 1+(1+\tau+z^*)\left(iL_{\varphi_1}+\overline{L}_{\varphi_1}L_{\varphi_2}+\overline{L}_{\varphi_2}L_{\varphi_1}+\mathsf{a}^*L_{\varphi_1}+L_{\varphi_2}\right)\|\mathsf{c}-\mathsf{c}'\|_2\\
&+(1+\tau)d_A(\mathsf{a},\mathsf{a}')+(1+\tau)L_{\varphi_2}\|\mathsf{c}-\mathsf{c}'\|_2\\
&\leq \text{máx}\{\|\mathsf{c}-\mathsf{c}'\|_2,d_A(\mathsf{a},\mathsf{a}')\}\left[1+(1+\tau+z^*)\left(iL_{\varphi_1}+\overline{L}_{\varphi_1}L_{\varphi_2}+\overline{L}_{\varphi_2}L_{\varphi_1}+\mathsf{a}^*L_{\varphi_1}+L_{\varphi_2}\right)\right.\\
&\left.+(1+\tau)(1+L_{\varphi_2})\right]\\
&\leq L_F\,\text{máx}\{\|\mathsf{c}-\mathsf{c}'\|_2,d_A(\mathsf{a},\mathsf{a}')\}.
\end{aligned}
$$

□

Ahora bien, sea ρ la densidad de la tasa de retorno ξ_t, de tal manera que la ley de transición se puede representar de la siguiente forma

$$K_{ij}^{\rho}(\mathsf{a},\mathsf{c})=\int_{\mathbb{R}}\mathbb{1}_j(F(i,\mathsf{c},\mathsf{a},z))\rho(z)\,dz. \tag{5.3}$$

Adicionalmente, para cada $i,j\in\mathcal{S}$ y $(\mathsf{a},\mathsf{c}),(\mathsf{a}',\mathsf{c}')\in A\times\mathbb{R}^d$ y como F toma valores en el conjunto $\mathcal{S}$, entonces la ley de transición K_{ij}^{ρ} será una función de Lipschitz, lo cual se observa a continuación:

$$\begin{aligned}\left|K_{ij}^{\rho}(\mathtt{a},\mathtt{c})-K_{ij}^{\rho}(\mathtt{a}',\mathtt{c}')\right| &= \left|\int_{\mathbb{R}}\mathbb{1}_j\left(F\left(i,\mathtt{c},\mathtt{a},z\right)\right)\rho\left(z\right)dz-\int_{\mathbb{R}}\mathbb{1}_j\left(F\left(i,\mathtt{c}',\mathtt{a}',z\right)\right)\rho\left(z\right)dz\right| \\ &= \left|\int_{\mathbb{R}}\left[\mathbb{1}_j\left(F\left(i,\mathtt{c},\mathtt{a},z\right)\right)-\mathbb{1}_j\left(F\left(i,\mathtt{c}',\mathtt{a}',z\right)\right)\right]\rho\left(z\right)dz\right| \\ &\leq \int_{\mathbb{R}}\left|\mathbb{1}_j\left(F\left(i,\mathtt{c},\mathtt{a},z\right)\right)-\mathbb{1}_j\left(F\left(i,\mathtt{c}',\mathtt{a}',z\right)\right)\right|\rho\left(z\right)dz \\ &\leq \int_{\mathbb{R}}\left|F\left(i,\mathtt{c},\mathtt{a},z\right)-F\left(i,\mathtt{c}',\mathtt{a}',z\right)\right|\rho\left(z\right)dz \\ &\leq L_F\,\text{máx}\left\{\left\|\mathtt{c}-\mathtt{c}'\right\|_2,d_A\left(\mathtt{a},\mathtt{a}'\right)\right\}\int_{\mathbb{R}}\rho\left(z\right)dz \\ &= L_F\,\text{máx}\left\{\left\|\mathtt{c}-\mathtt{c}'\right\|_2,d_A\left(\mathtt{a},\mathtt{a}'\right)\right\}\end{aligned}$$

Ahora procederemos a mostrar la ecuación en diferencias que determina la dinámica del sistema de proporciones $\left\{\mathtt{M}^N(t)\right\}$, como se realizó en la ecuación (2.3).

Dependiendo del sistema que estemos estudiando, podemos obtener una ecuación en diferencias para describir la evolución del proceso $\left\{\mathtt{M}^N(t)\right\}$; es decir, la ecuación descrita en la ecuación (2.6). En particular, para este ejemplo vamos a obtener una forma expliícita de la función G_ρ^N (2.6) en términos de variables aleatorias i.i.d. uniformemente distribuidas en $[0,1]$.

Este método consiste en simular el comportamiento evolutivo del sistema, para lo cual, consideraremos una partición del intervalo $[0,1]$ de la siguiente manera: dados $i,j\in\mathcal{S}$, $\mathtt{c}\in\mathbb{R}^d$ y $\mathtt{a}\in A$, definimos

$$\Delta_{ij}(\mathtt{a},\mathtt{c}) := \left[H_{ij}(\mathtt{a},\mathtt{c}),H_{i,j+1}(\mathtt{a},\mathtt{c})\right]\subseteq[0,1], \tag{5.4}$$

donde

$$H_{ij}(\mathtt{a},\mathtt{c}) := \sum_{l=1}^{j-1}K_{il}^{\rho}(\mathtt{a},\mathtt{c}) \quad \forall\, i,j\in\mathcal{S}. \tag{5.5}$$

Claramente, para cada $i\in\mathcal{S}$, el conjunto $\{\Delta_{ij}(\mathtt{a},\mathtt{c})\}$ es una partición del intervalo $[0,1]$; más aún, notemos que

$$\mathit{longitud}\left(\Delta_{ij}(\mathtt{a},\mathtt{c})\right) = K_{ij}^{\rho}(\mathtt{a},\mathtt{c}).$$

Así pues, a fin de poder representar la evolución de las proporciones $\{M_i^N(t)\}$ que corresponderán a los distintos niveles (clases) de riqueza, se necesita definir N variables aleatorias i.i.d., w_k^n en $[0,1]$ que se usarán más adelante. Finalmente, para cada $i\in\mathcal{S}$ y $t\in\mathbb{N}$, denotamos

$$\begin{aligned}\mathtt{w}^i(t) &:= \left(\mathtt{w}_1^i(t),\mathtt{w}_2^i(t),\ldots,\mathtt{w}^i_{N\mathtt{M}_i^N(t)}(t)\right), \\ \mathtt{w}_t &:= \left(\mathtt{w}^0(t),\mathtt{w}^1(t),\ldots,\mathtt{w}^S(t)\right).\end{aligned} \tag{5.6}$$

Cabe señalar que $\sum_{i=0}^{S}\mathtt{M}_i^N(t)\,N = N$, así que $\mathtt{w}_t\in[0,1]^N$. Este hecho implica que el número de variables aleatorias uniformes necesarias para la simulación de la evolución del proceso coincide con el número N de inversionistas, lo cual es consistente con lo expresado en la ecuación (2.6).

Por lo tanto, de acuerdo a lo señalado previamente, la proporción de objetos en el tiempo $t+1$ en la clase j estará dada por

$$\mathtt{M}_j^N(t+1) := \frac{1}{N}\sum_{i=0}^{S}\sum_{n=1}^{N\mathtt{M}_i^N(t)} \mathbb{1}_{\Delta_{ij}(\mathtt{a}(t),\mathtt{C}^N(t))}\left(w_n^i(t)\right). \tag{5.7}$$

Ahora si reescribimos dicha expresión en términos de la notación del Capítulo 2, tendremos que usar de igual manera la notación $\mathcal{Y}_N := \mathbb{P}_N(\mathcal{S}) \times \mathbb{R}^d$, $\mathcal{Y} := \mathbb{P}(\mathcal{S}) \times \mathbb{R}^d$ y $\mathtt{Y}^N(t) = \left(\mathtt{M}^N(t), \mathtt{C}^N(t)\right) \in \mathcal{Y}_N$. Así pues, para $i \in \mathcal{S}$, la función de evolución tomará la forma

$$G_{\rho,i}^N\left(\mathtt{Y}^N(t), \mathtt{a}(t), \mathtt{w}_t\right) := \frac{1}{N}\sum_{i=0}^{S}\sum_{n=1}^{N\mathtt{M}_i^N(t)} \mathbb{1}_{\Delta_{ij}(\mathtt{a}(t),\mathtt{C}^N(t))}\left(w_n^i(t)\right).$$

De lo anterior, se puede extraer una forma vectorial de la función G_ρ^N como sigue:

$$G_\rho^N(\mathtt{y}, a, \mathtt{w}) = \left(G_{\rho,0}^N(\mathtt{y}, a, \mathtt{w}), G_{\rho,1}^N(\mathtt{y}, a, \mathtt{w}), \ldots, G_{\rho,S}^N(\mathtt{y}, a, \mathtt{w})\right) \quad (\mathtt{y}, \mathtt{a}, \mathtt{w}) \in \mathcal{Y}_N \times A \times [0,1]^N. \tag{5.8}$$

Luego entonces, hemos obtenido la ecuación en diferencias que buscábamos; es decir,

$$\mathtt{M}^N(t+1) = G_\rho^N\left(\mathtt{M}^N(t), \mathtt{C}^N(t), \mathtt{a}(t), \mathtt{w}_t\right). \tag{5.9}$$

Lo que sigue es definir la dinámica del proceso $\left\{\mathtt{Y}^N(t)\right\}$. Para tal fin, supondremos que la transición del ambiente se rige por medio de una función $g : \mathbb{R}^d \times \mathbb{P}(\mathcal{S}) \times A \to \mathbb{R}$ arbitraria que cumple las condiciones del inciso (b) de la Hipótesis A. Así, el contexto del ambiente evoluciona de acuerdo con la siguiente ecuación

$$\mathtt{C}^N(t+1) = g\left(\mathtt{C}^N(t), \mathtt{M}^N(t+1), \mathtt{a}(t)\right) \quad t \in \mathbb{N}. \tag{5.10}$$

Juntando las piezas, al igual que se efectuó en la ecuación (2.7) y derivado de (5.9) y (5.10), se define la función

$$H_\rho^N(\mathtt{y}, a, \mathtt{w}) := \left(G_\rho^N(\mathtt{y}, a, \mathtt{w}), g\left(\mathtt{c}, G_\rho^N(\mathtt{y}, \mathtt{a}, \mathtt{w}), a\right)\right) \quad \forall\, (\mathtt{y}, a, \mathtt{w}) \in \mathcal{Y}_N \times A \times \mathbb{R}^N, \tag{5.11}$$

la cual determina la dinámica del proceso $\left\{\mathtt{Y}^N(t)\right\}$ al igual que se realizó en (2.8).

Finalmente, supongamos que las funciones de costo en cada etapa ℓ, así como el costo terminal ℓ_T son acotadas por una constante $R > 0$ y uniformemente de Lipschitz con constante L_l, estos representan los gastos del gobierno o de la entidad reguladora por aplicar un subsidio o un impuesto en cada etapa del tiempo. Esto es

$$|\ell(\mathtt{y}, a)| + |\ell_T(\mathtt{y})| \leq R \quad \forall a \in A,\ \mathtt{y} \in \mathcal{Y}_N,$$

$$\sup_{(a,a') \in A \times A} |\ell(\mathtt{y}, a) - \ell(\mathtt{y}', a')| + |\ell_T(\mathtt{y}) - \ell_T(\mathtt{y}')| \leq L_l \|\mathtt{y} - \mathtt{y}'\| \quad \forall\, \mathtt{y}, \mathtt{y}' \in \mathcal{Y}_N, a, a' \in A.$$

Por otra parte, dado que el espacio de acciones A es finito, la continuidad de $a \to H_{\rho}^{N}(\cdot, a, \cdot)$ se sigue de inmediato, por lo que el inciso (c) de la Hipótesis A se valida automáticamente. Entonces podemos concluir que nuestro ejemplo satisface lo requerido en la Hipótesis A.

Cabe señalar que, a partir de las funciones de costo por etapa ℓ así como del costo terminal ℓ_T y de forma similar que se realizó en el Capítulo 2, podemos definir el costo total esperado de una política $\pi^N \in \Pi^N$, mismo que estará dado por

$$V^N\left(\pi^N, \mathtt{y}\right) := E_{\mathtt{y}}^{\pi^N}\left[\sum_{t=0}^{T-1} \ell\left(\mathtt{Y}^N(t), \mathtt{a}(t)\right) + \ell_T\left(\mathtt{Y}^N(T)\right)\right], \tag{5.12}$$

donde, al igual que antes, la esperanza se considera con respecto a la medida de probabilidad $P_{\mathtt{y}}^{\pi^N}$ inducida por la política π^N y el estado inicial $\mathtt{Y}^N(0) = \mathtt{y}$.

Por lo que, el problema de control óptimo es determinar una política $\pi_*^N \in \Pi^N$ generada por los subsidios o impuestos, de tal forma que:

$$V^N\left(\pi_*^N, \mathtt{y}\right) = \inf_{\pi^N \in \Pi^N} V^N\left(\pi^N, \mathtt{y}\right) =: V_*^N(\mathtt{y}) \qquad \forall\, \mathtt{y} \in \mathcal{Y}_N.$$

Por otra parte, observemos que los hechos establecido hasta este punto, se ha demostrado que se satisface lo requerido en la Hipótesis A, esto quiere decir que, para encontrar el óptimo del modelo en cuestión, basta analizar el modelo de control de campo medio asociado a éste.

Así pues, definimos el modelo de control de campo medio (determinístico) $\{\mathtt{M}(t), \mathtt{C}(t)\} \in \mathcal{Y}$ como sigue

$$\mathtt{M}(t+1) = \mathtt{M}(t)\, K_\rho(\mathtt{a}(t), \mathtt{C}(t)), \tag{5.13}$$

$$\mathtt{C}(t+1) = g(\mathtt{M}(t+1), \mathtt{C}(t), \mathtt{a}(t)), \tag{5.14}$$

donde K_ρ es la matriz $\left(K_{ij}^{\rho}\right)$, donde los elementos de esta misma se encuentran definidas en (5.3) y $g : \mathbb{R}^d \times \mathbb{P}(\mathcal{S}) \times A \to \mathbb{R}$ es la función definida en (5.10). Ahora bien, observemos que los componentes del vector $\mathtt{M}(t+1)$ están dados por

$$\mathtt{M}_j(t+1) = \sum_{i=0}^{S} \mathtt{M}_i(t)\, K_{ij}^{\rho}(\mathtt{a}(t), \mathtt{c}(t)). \tag{5.15}$$

Donde $\mathtt{M}(0) = \mathtt{m}_0 \in \mathbb{P}(\mathcal{S})$. Además la función G_ρ en (3.1) queda como

$$G_\rho(\mathtt{m}, \mathtt{c}, a) = \mathtt{m} K_\rho(a, \mathtt{c}) \qquad (\mathtt{m}, \mathtt{c}) \in \mathcal{Y}, a \in A. \tag{5.16}$$

Es claro que G_ρ es Lipschitz ya que K_{ij}^{ρ} lo es. Denotemos como L_G su constante de Lipschitz asociada.

Una vez definidos tanto las proporciones como el contexto del ambiente en el modelo de control de campo medio, podemos definir la dinámica del mismo. Para ello, consideremos a $H_\rho : \mathcal{Y} \times A \to \mathcal{Y}$ la cual tiene la siguiente estructura:

$$H_\rho(y, a) := (G_\rho(y, a), g(\mathtt{c}, G_\rho(y, a), a)) \quad \forall\, y = (m, c) \in \mathcal{Y},\ a \in A. \tag{5.17}$$

Y al igual en el Capítulo 3, podemos escribir la función de la dinámica del proceso $\{(\mathtt{M}(t), \mathtt{C}(t))\}$, de la siguiente manera:

$$\begin{aligned}\mathtt{Y}(t+1) &= (G_\rho(\mathtt{Y}(t), \mathtt{a}(t)), g(\mathtt{Y}(t), \mathtt{a}(t))) \\ &= (G_\rho(\mathtt{Y}(t), \mathtt{a}(t)), g(G_\rho(\mathtt{Y}(t), \mathtt{a}(t)), \mathtt{C}(t), \mathtt{a}(t))) \\ &= H_\rho(\mathtt{Y}(t), \mathtt{a}(t)),\end{aligned} \tag{5.18}$$

con $\mathtt{Y}(0) = (\mathtt{m}_0, \mathtt{c}_0) \in \mathcal{Y}$.

El problema de control asociado al modelo de campo medio, se define exactamente como en (3.7)-(3.8), junto con sus ecuaciones de programación dinámica como en (3.10)-(3.11).

Para concluir el análisis, demostraremos que las trayectorias definidas en (5.18) asociadas al modelo de campo medio (modelo límite), son justamente trayectorias límite de aquellas dadas en (5.9)-(5.10). Para este propósito, demostraremos la validez del inciso (b) de la Hipótesis B.

En lo que sigue usaremos la notación establecida en el Capítulo 4. Sea $\pi = \{f_t\} \in \Pi$ una política arbitraria y $\mathtt{y} \in \mathcal{Y}_N \subseteq \mathcal{Y}$ el estado inicial del proceso. Para lo que, recordaremos la dinámica del proceso con N inversionistas

$$\mathtt{M}_j^N(t+1) := \frac{1}{N} \sum_{i=0}^{S} \sum_{n=1}^{N\mathtt{M}_i^N(t)} \mathbb{1}_{\Delta_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right)}\left(w_n^i(t)\right), \tag{5.19}$$

y a fin de simplificar la notación definimos

$$B_{inj}^N := \mathbb{1}_{\Delta_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right)}\left(w_n^i(t)\right) \qquad i, j \in \mathcal{S}, n \in \mathbb{N}. \tag{5.20}$$

Notemos que, para cada $t \in \mathbb{N}$, $\left\{B_{inj}^N(t)\right\}_{inj}$ es una familia de variables aleatorias independientes e identicamente distribuidas con distribución de Bernoulli, asimismo, notemos que la esperanza de estas mismas está dada por

$$\mathbb{E}_{\mathtt{y}}^{\pi}\left[B_{inj}^N(t) \,|\mathtt{a}^{\pi^N}(t) = \mathtt{a}, \mathtt{C}^N(t) = \mathtt{c}\right] = K_{ij}^{\rho}(\mathtt{a}, \mathtt{c}).$$

Ahora bien, de acuerdo con lo anterior notemos que la ecuación (5.19) queda ahora como sigue

$$\mathtt{M}_j^N(t+1) := \frac{1}{N} \sum_{i=0}^{S} \sum_{n=1}^{N\mathtt{M}_i^N(t)} B_{inj}^N(t). \tag{5.21}$$

Y a partir de esta mismas y por las ecuaciones (2.1) y (2.3) de [11], se puede obtener lo siguiente

$$\mathbb{P}_{\mathtt{y}}^{\pi}\left\{\left|\sum_{n=1}^{N\mathtt{M}_i^N(t)} B_{inj}^N(t) - N\mathtt{M}_i^N(t) K_{ij}^{\rho}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right)\right| \geq N\varepsilon\right\} \leq 2e^{-2N\varepsilon^2}.$$

Entonces, debido a esto y a una simple propiedad de la probabilidad se sigue

$$\mathbb{P}^{\pi}_{\mathtt{y}}\left\{\left|\sum_{n=1}^{N\mathtt{M}_i^N(t)} B^N_{inj}(t) - N\mathtt{M}_i^N(t)\, K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right)\right| < N\varepsilon\right\} > 1 - 2e^{-2N\varepsilon^2}.$$

Asimismo, consideremos el siguiente conjunto

$$\overline{\Omega} = \left\{\omega \in \Omega \ : \left|\sum_{n=1}^{N\mathtt{M}_i^N(t)} B^N_{inj}(t) - N\mathtt{M}_i^N(t)\, K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right)\right| < N\varepsilon\right\} \subseteq \Omega. \tag{5.22}$$

Adicionalmente, para $t \in \mathbb{N}$, sea $\varepsilon_t > 0$ un número tal que

$$\left\|\mathtt{Y}^N(t) - \mathtt{Y}(t)\right\| \leq \varepsilon_t, \tag{5.23}$$

por lo que, en particular obtenemos para los procesos de proporciones y del contexto del ambiente

$$\left\|\mathtt{M}^N(t) - \mathtt{M}(t)\right\|_1 \leq \varepsilon_t \qquad \left\|\mathtt{C}^N(t) - \mathtt{C}(t)\right\|_2 \leq \varepsilon_t. \tag{5.24}$$

Por lo que, de las ecuaciones (5.13), (5.19) y (5.24) y considerando el conjunto $\overline{\Omega}$ tenemos

$$\begin{aligned}
\left|\mathtt{M}_j^N(t+1)\right. &\left. -\mathtt{M}_j(t+1)\right| = \left|\frac{1}{N}\sum_{i=0}^{S}\sum_{n=1}^{N\mathtt{M}_i^N(t)} B^N_{inj}(t) - \sum_{i=0}^{S}\mathtt{M}_i(t)\, K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}(t)\right)\right| \\
&\leq \left|\frac{1}{N}\sum_{i=0}^{S}\left[\sum_{n=1}^{N\mathtt{M}_i^N(t)} B^N_{inj}(t) - N\mathtt{M}_i(t)\, K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}(t)\right)\right]\right| \\
&\leq \frac{1}{N}\sum_{i=0}^{S}\left|\sum_{n=1}^{N\mathtt{M}_i^N(t)} B^N_{inj}(t) - N\mathtt{M}_i(t)\, K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}(t)\right)\right| \\
&\leq \frac{1}{N}\sum_{i=0}^{S}\left|\sum_{n=1}^{N\mathtt{M}_i^N(t)} B^N_{inj}(t) - N\mathtt{M}_i^N(t)\, K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right)\right| \\
&+ \sum_{i=0}^{S}\left|\mathtt{M}_i^N(t) - \mathtt{M}_i(t)\right| K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right) \\
&+ \sum_{i=0}^{S}\mathtt{M}_i(t)\left|K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right) - K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}(t)\right)\right| \\
&\leq \frac{1}{N}\sum_{i=0}^{S} N\varepsilon + \sum_{i=0}^{S}\varepsilon_t K^{\rho}_{ij}\left(\mathtt{a}^{\pi^N}(t), \mathtt{C}^N(t)\right) + \sum_{i=0}^{S} L_K \mathtt{M}_i(t)\left\|\mathtt{C}^N(t) - \mathtt{C}(t)\right\|_2 \\
&\leq (S+1)\varepsilon + (S+1)\varepsilon_t + L_K\varepsilon_t.
\end{aligned} \tag{5.25}$$

Y debido a que en la desigualdad anterior no hay dependencia sobre $j \in \mathcal{S}$ se obtiene lo siguiente

$$\left\|\mathtt{M}^N(t+1) - \mathtt{M}(t+1)\right\|_1 \leq (S+1)\varepsilon + (S+1)\varepsilon_t + L_K\varepsilon_t.$$

Ahora nos centraremos en acotar el ambiente del contexto y dado que g es una función de Lipschitz y de las ecuaciones (5.10), (5.14), (5.24) y (5.25) resulta

$$\begin{aligned}
&\left\|\mathtt{C}^N(t+1)-\mathtt{C}(t+1)\right\|_2=\left\|g\left(\mathtt{C}^N(t),\mathtt{M}^N(t+1),\mathtt{a}^{\pi^N}(t)\right)-g\left(\mathtt{C}(t),\mathtt{M}(t+1),\mathtt{a}^{\pi^N}(t)\right)\right\|_2\\
&\leq L\,\text{máx}\left\{\left\|\mathtt{M}^N(t+1)-\mathtt{M}(t+1)\right\|_1,\left\|\mathtt{C}^N(t)-\mathtt{C}(t)\right\|_2,d_A\left(\mathtt{a}^{\pi^N}(t),\mathtt{a}^{\pi^N}(t)\right)\right\}\\
&\leq L\,\text{máx}\left\{(S+1)\varepsilon+(S+1)\varepsilon_t+L_K\varepsilon_t,\varepsilon_t\right\}=L\left((S+1)\varepsilon+(S+1)\varepsilon_t+L_K\varepsilon_t\right).
\end{aligned}$$

Entonces a paritr de los resultados obtenidos podemos acotar las diferencias entre los procesos $\left\{\mathtt{Y}^N(t)\right\}$ y $\{\mathtt{Y}(t)\}$ como sigue

$$\begin{aligned}
\left\|\mathtt{Y}^N(t+1)-\mathtt{Y}(t+1)\right\|&=\text{máx}\left\{\left\|\mathtt{M}^N(t+1)-\mathtt{M}(t+1)\right\|_1,\left\|\mathtt{C}^N(t+1)-\mathtt{C}(t+1)\right\|_2\right\}\\
&<\text{máx}\left\{(S+1)\varepsilon+(S+1)\varepsilon_t+L_K\varepsilon_t,L\left((S+1)\varepsilon+(S+1)\varepsilon_t+L_K\varepsilon_t\right)\right\}\\
&=L\left((S+1)\varepsilon+(S+1)\varepsilon_t+L_K\varepsilon_t\right).
\end{aligned}$$

Y considerando el caso para $t=0$, se sigue

$$\left\|\mathtt{Y}^N(0)-\mathtt{Y}(0)\right\|=\varepsilon_0=0.$$

Y aplicando un proceso de inducción sobre la variable, nos lleva a concluir que

$$\left\|\mathtt{Y}^N(t+1)-\mathtt{Y}(t+1)\right\|<L(S+1)\varepsilon\beta_t.$$

Donde $\{\beta_t\}$ es una sucesión creciente. Entonces, si fijamos un $T\in\mathbb{N}$ se obtiene

$$\forall 0\leq t\leq T\quad\left\|\mathtt{Y}^N(t+1)-\mathtt{Y}(t+1)\right\|<L(S+1)\varepsilon\beta_T.$$

Y debido a que esto es válido para el conjunto $\overline{\Omega}$, podemos asumir que bajo la política $\pi\in\Pi$ se satisface

$$\mathbb{P}_{\mathtt{y}}^{\pi}\left\{\sup_{o\leq t\leq T}\left\|\mathtt{Y}^N(t+1)-\mathtt{Y}(t+1)\right\|<L(S+1)\varepsilon\beta_T\right\}>1-2e^{-2N\varepsilon^2}.$$

Ahora bien, si definimos $\gamma_T(\varepsilon):=L(S+1)\varepsilon\beta_T, K=\lambda=2$ y de las propiedades de una medida de probabilidad

$$\sup_{\pi\in\Pi}\mathbb{P}_{\mathtt{y}}^{\pi}\left\{\sup_{0\leq t\leq T}\left\|\mathtt{Y}^N(t)-\mathtt{Y}(t)\right\|\geq\gamma_T(\varepsilon)\right\}\leq KTe^{-\lambda N\varepsilon^2}. \tag{5.26}$$

Y observemos que $\gamma_T(\varepsilon)\to 0$ cuando $\varepsilon\to 0$, por lo que, se ha demostrado lo deseado, más aún, este mismo resultado coincide con el inciso (b) de la Hipótesis B.

A manera de conclusión, podemos decir que para determinar el óptimo del problema con los N inversionistas, basta considerar el modelo deterministico de campo medio y obtener políticas óptimas —estas últimas obtenidas de la ecuación de programación dinámica (3.10)–(3.11)— para luego aplicarlas al modelo original. El alcance de este ejemplo no da para resolver explícitamente tales políticas; sin embargo, asegura que el problema es soluble.

5.2 Problema de intermediación en redes

Supongamos que existen I clientes, los cuales mandan solicitudes (tareas) a un dispositivo central, el cual, a su vez, asigna y direcciona las tareas recibidas a d servidores, cada uno con un número de procesadores fijos. La idea general de este problema es minimizar el tiempo de espera de atención de las solicitudes. Cabe señalar que en cada etapa del tiempo se puede monitorear si cada uno de estos clientes se encuentra conectado o desconectado; así que, para cada $j \in \mathcal{I} := \{1, 2, \ldots, I\}$, denotaremos por

$$\begin{array}{ll} Y_j(t) = 1 & \text{si el cliente } j \text{ se encuentra conectado en el tiempo } t, \\ Y_j(t) = 0 & \text{si el cliente } j \text{ se encuentra desconectado en el tiempo } t. \end{array}$$

El modelo supondrá que si el cliente se encuentra conectado, éste manda una sola solicitud entre t y $t+1$. De lo anterior, el número de solicitudes totales enviadas al tiempo t es

$$Y(t) := \sum_{j=1}^{I} Y_j(t).$$

Por otra parte, cada uno de los $i \in \{1, \cdots, d\}$ servidores, cuenta con P_i procesadores, respectivamente, y al igual que en el caso de los clientes, cada uno de estos procesadores se puede encontrar disponible o fuera de servicio, según sea el caso, entonces de forma análoga denotamos a

$$\begin{array}{ll} X_{ij}(t) = 1 & \text{si el procesador } j \text{ del servidor } i \text{ se encuentra disponible en el tiempo } t, \\ X_{ij}(t) = 0 & \text{si el procesador } j \text{ del servidor } i \text{ se encuentra fuera de servicio en el tiempo } t. \end{array}$$

Adicionalmente, cada uno de estos servidores atiende las solicitudes que recibe a una velocidad uniforme μ_i, siempre y cuando se encuentre disponible.

Así que, el número total de procesadores disponibles en el servidor i en el periodo de tiempo comprendido entre t y $t+1$ está dado por

$$X_i(t) := \sum_{j=1}^{P_i} X_{ij}(t).$$

Definimos el número de solicitudes en espera de ser atendidas en el servidor i como $B_i(t)$. De esta manera, para cada $t \in \mathbb{N}$, el estado del sistema evoluciona debido al envío de las $Y(t)$ solicitudes generadas por los clientes, mismas que son asignadas a cada uno de los servidores mediante el dispositivo central, el cual elige una acción $\mathsf{a}(t) \in [0, 1]$ a través de la probabilidad de que una solicitud se envíe al servidor i en el tiempo t. Adicionalmente el número de solicitudes en espera en el servidor i evoluciona de acuerdo a lo siguiente

$$B_i(t+1) = \left(B_i(t) - \mu_i X_i(t) + \mathsf{a}_i Y(t)\right)^+.$$

A partir de este mismo, podemos definir el vector $B(t)$, el cual denotará el número de solicitudes en espera en cada uno de los servidores de todo el sistema, es decir:

$$B(t) = (B_1(t), B_2(t), \ldots, B_d(t)).$$

A continuación definimos el costo por etapa en el tiempo t, así como el costo terminal $T \in \mathbb{N}$, como el número de solicitudes en espera de ser atendidas, es decir,

$$\ell(t) = \sum_{i=1}^{d} B_i(t), \quad \ell_T(t) = \sum_{i=1}^{d} B_i(T),$$

Una vez definidos todos los elementos previos, podemos formular este problema como un problema de interacción de partículas con un ambiente en común, de acuerdo con los siguientes puntos.

1. Las $N := I + \sum_{i=1}^{d} P_i$ partículas, cada una de estas puede ser un cliente (denotado por s) o bien un procesador (perteneciente a su respectivo servidor i) y además puede estar conectado (disponible), denotado por 1, o desconectado (fuera de servicio), denotado por 0, según corresponda. Proponemos como estados del sistema al conjunto $\mathcal{S}$, definido como sigue

$$\mathcal{S} := \{(x, e) \ : \ x \in \{s, 1, 2, \ldots, d\}, e \in \{0, 1\}\}.$$

2. La proporción $\mathtt{M}^N(t)$ se define como la proporción de clientes conectados y las proporciones de procesadores disponibles para cada uno de los servidores. Notemos que para cada $(x, e) \in \mathcal{S}$ su proporción se ve como

$$\begin{aligned} \mathtt{M}^N_{(x,e)}(t) &= \frac{1}{N} Y(t) \qquad \text{si } x = s \\ \mathtt{M}^N_{(x,e)}(t) &= \frac{1}{N} X_x(t) \qquad \text{si } x = 1, 2, \ldots, d \end{aligned}$$

3. Como se dijo anteriormente, las acciones del dispositivo central serán las elecciones de dirección de cada una de las solicitudes, es decir, $\mathtt{a}_i(t)$ es la probabilidad de que una solicitud sea enviada al servidor i en el tiempo t.

4. El ambiente del contexto del sistema en el tiempo t nos describirá las solicitudes que se tienen en espera en cada uno de los servidores a dicho tiempo, por lo que, este mismo dependerá del vector $B(t)$, mismo que definimos como

$$\mathtt{C}^N(t) := \left(\frac{B_1(t)}{d}, \frac{B_2(t)}{d}, \ldots, \frac{B_d(t)}{d} \right).$$

5. El kernel de transición que describirá la mecánica de evolución del sistema será independiente de la acción tomada por el dispositivo central, así como de las solicitudes en espera, más aún, la probabilidad de ir del estado $(x, e) \in \mathcal{S}$ al estado $(y, f) \in \mathcal{S}$ será 0 si $x \neq y$, debido a que un cliente no se puede convertir en un servidor y viceversa, es decir, $K^{\rho}_{(x,e),(y,f)}(\mathtt{a}, \mathtt{c}) = 0$, para los otros casos, supondremos que dicha probabilidad está dada.

Notemos que de lo descrito previamente y del numeral 4, podemos definir lo siguiente

$$B_i(t+1) = g_i\left(\mathtt{M}^N(t+1), \mathtt{C}^N(t), \mathtt{a}(t)\right) := \left(B_i(t) - \mu_i X_i(t) + \mathtt{a}_i Y(t)\right)^+,$$

por lo que, en general tendremos la siguiente expresión para la evolución del contexto del ambiente

$$\mathtt{C}^N(t) = g\left(\mathtt{M}^N(t+1), \mathtt{C}^N(t), \mathtt{a}(t)\right),$$

donde la función g es aquella que, tiene como elementos a cada una de las funciones g_i. Así que, a partir de los numerales anteriores se puede definir el modelo de control de campo medio como sigue

$$\mathtt{M}(t+1) = \mathtt{M}(t)\, K_\rho(\mathtt{a}(t), \mathtt{C}(t)), \tag{5.27}$$

$$\mathtt{C}(t+1) = g(\mathtt{M}(t+1), \mathtt{C}(t), \mathtt{a}(t)), \tag{5.28}$$

donde K_ρ es la matriz $\left(K^\rho_{(x,e),(y,f)}\right)$, donde los elementos de esta misma se encuentran definidas en el numeral 5 anterior, y g es la función definida previamente.

Ahora bien, debido al planteamiento del problema y siguiendo un proceso similar al efectuado en la referencia [6, Capítulo 4], se puede demostrar que el modelo planteado de esta manera, cumple lo requerido en la Hipótesis A. Además, simulando exactamente igual al proceso de proporciones M^N como se hizo en (5.19)-(5.20), se puede verificar que la Hipótesis B también es válida.

Por lo tanto, en virtud de lo obtenido en capítulos anteriores, se puede demostrar la existencia de una política óptima de asignación de tareas, pero asociada al modelo de campo medio (5.27) - (5.28), misma que es obtenida por medio de las ecuaciones de programación dinámica (3.10)-(3.11). Dicha política es la que aproxima de manera satisfactoria el mínimo tiempo de espera de solicitudes asociadas al problema original.

Bibliografía

[1] Achdou, I. and Capuzzo-Dolcetta, I. (2010). Mean field games: numerical methods. *SIAM J. Numer. Anal.*, 48:1136–1162.

[2] Bäuerle, N. (2021). Mean-field Markov decision processes. *Preprint*, pages 1–22.

[3] Bensoussan, A., Frehse, J., and Yam, P. (2013). *Mean field games and mean field control theory*. Springer Briefs in Mathematics, New York.

[4] Gast, N. and Gaujal, B. (2009). A mean field approach for optimization in particle systems and applications. *Institut National De Recherche En Informatiqueet et en Automatique*.

[5] Gast, N. and Gaujal, B. (2011). A mean field approach for optimization in discrete time. *Discrete Event Dyn. Syst*, 21:63–101.

[6] Gast, N., Gaujal, B., and Le Boudec, J. (2012). Mean field for Markov decision processes: from discrete to continuous optimization. *IEEE Trans. Autom. Control*, 57:2266–2280.

[7] Gomes, D., Mohr, J., and Souza, R. (2010). Discrete time finite state space mean field games. *J. Math. Pures Appl.*, 93:308–328.

[8] Hernández-Lerma, O. and Lasserre, J. (1996). *Discrete-time Markov control processes*. Springer-Verlag, New York.

[9] Higuera-Chan, C., Jasso-Fuentes, H., and Minjárez-Sosa, J. (2016). Discrete-time control for systems of interacting objects with unknown random disturbance distributions: a mean field approach. *Appl. Math. Optim*, 74:197–227.

[10] Higuera-Chan, C., Jasso-Fuentes, H., and Minjárez-Sosa, J. (2017). Control systems of interacting objects modeled as a game against nature under a mean field approach. *Journal of Dynamics and Games*, 4:59–74.

[11] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–20.

[12] Huang, M. (2010). Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM J. Control Optim.*, 48:3318–3353.

[13] Huang, M., Palhamé, R., and Caines, P. (2006). Large-population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6:221–256.

[14] Kolokoltsov, V., Troeva, N., and Yang, W. (2014). On the rate of convergence for the mean-field approximation of controlled diffusions with large number of players. *Dyn. Games Appl.*, 4:208–230.

[15] Lasry, J. and Lions, P. (2007). Mean field games. *Jap. J. Math.*, 2:229–260.

[16] Martínez-Manzanares, M. and Minjárez-Sosa, J. (2008). A mean-field absorbing control model for interacting object systems. *Preprint*, pages 1–27.

[17] Saldi, N., Basar, T., and Raginsky, M. (2018). Markov-Nash equilibria in mean-field games with discounted cost. *SIAM J. Control Optim.*, 56:4256–4287.

[18] Wiecek, P. (2020). Discrete-time ergodic mean-field games with average reward on compact spaces. *Dynamic Games and Applications*, 10:222–256.