

# SOBRE REDES NEURONALES Y CONTROL ÓPTIMO

T E S I S

Que para obtener el grado de

**Maestro en Ciencias**

con especialidad en

**Probabilidad y Estadística**

**Presenta**

Erik Alejandro Gallegos Baños

**Director de Tesis:**

Daniel Hernández Hernández

---

Autorización de la versión final





Centro de Investigación en Matemáticas, A.C.



# ACTA PROVISIONAL

## Acta de Examen de Grado

Acta No.: 161

Libro No.: 003

Foja No.: 161

En la Ciudad de Guanajuato, Gto., siendo las 11:00 horas del día 21 de agosto del año 2020, se reunieron los miembros del jurado integrado por los señores:

**DR. MARIANO JOSÉ JUAN RIVERA MERAZ** (CIMAT)  
**DR. JOSÉ LUIS ÁNGEL PÉREZ GARMENDIA** (CIMAT)  
**DR. DANIEL HERNÁNDEZ HERNÁNDEZ** (CIMAT)

Bajo la presidencia del primero y con carácter de secretario el segundo, para proceder a efectuar el examen que para obtener el grado de



**MAESTRO EN CIENCIAS  
CON ESPECIALIDAD EN PROBABILIDAD Y ESTADÍSTICA**

Sustenta

**ERIK ALEJANDRO GALLEGOS BAÑOS**

En cumplimiento con lo establecido en los reglamentos y lineamientos de estudios de posgrado del Centro de Investigación en Matemáticas, A.C., mediante la presentación de la tesis

**“SOBRE REDES NEURONALES Y CONTROL ÓPTIMO”**


Los miembros del jurado examinaron alternadamente al (la) sustentante y después de deliberar entre sí resolvieron declararlo (a)

**APROBADO**



**CIMAT**  
DIRECCIÓN  
GENERAL

Dr. Victor Manuel Rivero Mercado  
Director General

  
\_\_\_\_\_  
**DR. MARIANO JOSÉ JUAN RIVERA MERAZ**  
Presidente

  
\_\_\_\_\_  
**DR. JOSÉ LUIS ÁNGEL PÉREZ GARMENDIA**  
Secretario

  
\_\_\_\_\_  
**DR. DANIEL HERNÁNDEZ HERNÁNDEZ**  
Vocal

000000

ACTA PROVISIONAL



*A mi familia.*



# Agradecimientos

Agradezco a CONACyT por el apoyo económico provisto mediante la asignación de una beca de estudios de posgrado. A CIMAT por las facilidades brindadas para mi desarrollo durante el período en que estudié la maestría en probabilidad y estadística. También a mi asesor, el doctor Daniel Hernández, por el tiempo y guía ofrecida para desarrollar el presente trabajo.





# Resumen

El presente trabajo es de carácter expositivo. Su objetivo global es servir como texto de partida para matemáticos que quieran colaborar en la formulación de un marco teórico para entender el problema de entrenamiento de una red neuronal.

En el capítulo 1 se recaba el vocabulario usado por la comunidad de aprendizaje máquina para describir el problema de entrenamiento.

En el capítulo 2, siguiendo [E et al., 2018], se presenta el problema de entrenamiento de una red neuronal residual como la discretización de un problema de control, referido como problema empírico. Éste a su vez surge de otro problema de control, llamado problema teórico, al tomar una muestra iid. Para el problema teórico se prueba un Principio del Máximo de Pontryagin.

En el siguiente capítulo se estudia la relación del control óptimo para el problema empírico y el del problema teórico. Ahí se demuestra que bajo cierta suposición, el control óptimo del problema empírico es consistente, es decir, converge en probabilidad al control óptimo del problema teórico.

Finalmente en el capítulo 4 se describe el algoritmo EMSA, un algoritmo de entrenamiento de redes residuales basado en el Principio del Máximo de Pontryagin.

## Palabras clave

Redes Residuales, Principio del Máximo de Pontryagin, Algoritmo EMSA.



# Índice general

<b>Agradecimientos</b>	<b>I</b>
<b>Resumen</b>	<b>II</b>
<b>Introducción</b>	<b>1</b>
<b>1. Preliminares</b>	<b>3</b>
1.1. Preliminares computacionales . . . . .	3
1.2. Preliminares matemáticos . . . . .	6
<b>2. Problema de entrenamiento como problema de control</b>	<b>11</b>
2.1. Formulación matemática . . . . .	11
2.2. Principio de Programación Dinámica . . . . .	13
2.3. Principio del Máximo de Pontryagin . . . . .	19
<b>3. Relación entre controles óptimos de problemas empírico y teórico</b>	<b>23</b>
3.1. Aproximaciones aleatorias a ceros de transformaciones de espacios de Banach.	24
3.2. Verificación de condiciones . . . . .	28
<b>4. Algoritmo EMSA</b>	<b>39</b>
4.1. Algoritmo EMSA y sus propiedades . . . . .	40
<b>Conclusiones y Trabajo Futuro</b>	<b>47</b>
Conclusiones . . . . .	49
Trabajo Futuro . . . . .	50
<b>Bibliografía</b>	<b>50</b>



# Introducción

En los últimos años, redes neuronales han mostrado ser un método muy eficaz para el problema de clasificación en aprendizaje máquina. Este problema puede interpretarse como aproximar una función desconocida  $F : X \rightarrow Y$ , que en principio es muy complicada, partiendo de una muestra  $\{(x^i, y^i)\}_{i=1}^N$ . La red neuronal representa una forma de construir una función que aproxima  $F$ . Esta aproximación es de la forma  $x_T(\cdot, \theta)$  donde  $\theta \in \mathbb{R}^m$  es un parámetro a encontrar y  $T$  es un parámetro del diseño de la aproximación. Entrenar la red neuronal significa encontrar un  $\theta$  tal que  $x_T(\cdot, \theta)$  aproxime bien a  $F$  en general, y no sólo en los valores  $x^i$  de la muestra. En su mayoría, la forma de diseñar la red, es decir cómo construir  $x_T$ , ha sido experimental; además no se puede paralelizar la parte de optimización que construye o encuentra el parámetro  $\theta$ . Más aún, no hay un marco teórico que dé garantías sobre la eficacia del parámetro  $\theta$  encontrado al entrenar la red, por decir, que se asegure en algún sentido que es cercano al ideal, o que no cambia mucho si se cambia la muestra de entrenamiento. Por otro lado, el vocabulario particular del área de aprendizaje máquina suele ser muy poco conocido entre la comunidad matemática. Esto ha limitado la colaboración de ambos grupos para establecer un marco teórico que aborde las carencias antes mencionadas.

En 2017 un grupo de investigadores, considerando un tipo particular de redes llamadas redes residuales, interpretaron el problema de entrenamiento como una discretización de un problema de control óptimo. Suponiendo que se cumplía el Principio del Máximo de Pontryagin para dicho problema teórico, propusieron un nuevo método para entrenar redes residuales. Mediante experimentos numéricos observaron que su propuesta requería menos iteraciones para entrenar, además de que se podía paralelizar. Un año después, en 2018, algunos de estos científicos se dedicaron a estudiar las propiedades matemáticas de dicho problema teórico, y publicaron un artículo donde enunciaban varios resultados, imitando la teoría para control clásico. Desafortunadamente casi todas las demostraciones, del artículo de 2018, contienen errores o deficiencias de algún tipo que dificultan determinar la veracidad del argumento y proposición.

Así, partiendo de lo anterior, este trabajo tiene tres objetivos. Por una parte es rescatar las proposiciones más importantes del artículo donde se estudian las propiedades matemáticas del problema teórico, por otra parte escribir un texto que sea fácil de leer para la comunidad matemática, esperando que sirva para que más personas puedan colaborar en el área del marco teórico de redes neuronales. Finalmente, presentar la clase de resultados que se buscan desde el punto de vista computacional en lo que respecta a algoritmos de entrenamiento de redes neuronales.



# Capítulo 1

## Preliminares

Este capítulo consiste de dos secciones, contiene vocabulario y resultados usados en el resto del trabajo; su principal función es como apartado de referencia. En preliminares computacionales, sección 1.1, se introduce terminología usada dentro de la comunidad de aprendizaje profundo. Ésta es empleada en su mayoría sólo hasta el capítulo 4. La sección de preliminares matemáticos, 1.2, contiene definiciones y resultados auxiliares usados en el trabajo, principalmente en el capítulo 2.

La sección 1.1 está basada en los libros [Goodfellow et al., 2016] y [Zhang et al., 2020]. La información para la sección 1.2 se basa principalmente en [Cardaliaguet, 2013], [Bressan and Piccoli, 2007] y [Ambrosio et al., 2005].

### 1.1. Preliminares computacionales

Desde hace varios años se ha intentado que las máquinas sean inteligentes. Una perspectiva de interpretar y abordar el problema propone lograr que una máquina sea capaz de realizar una tarea específica sin que se le proporcione las instrucciones o conocimiento humano de cómo hacerlo. Esta propuesta originó lo que hoy en día se conoce como **aprendizaje máquina**. Bajo el enfoque de éste, una máquina o **algoritmo de aprendizaje** es entrenado en vez de ser explícitamente programado. Un algoritmo de aprendizaje se dice **supervisado** si para entrenarlo se le debe proporcionar una colección de ejemplos de la encomienda en cuestión. Después de examinar la muestra, el algoritmo desarrollará una manera de hacer la tarea. Por ejemplo, si el objetivo es, dada una imagen con una prenda de vestir identificar el tipo de vestimenta en la imagen, entonces a un algoritmo de aprendizaje supervisado se le tendría que proporcionar una muestra de imágenes, cada una con una prenda, y la etiqueta del tipo de prenda de cada imagen. Después de entrenar el algoritmo, éste deberá ser capaz de determinar la vestimenta en una nueva imagen que se le proporcione.

Para fines del presente trabajo, se entenderá una **red neuronal**, o NN por su nombre en inglés, como un algoritmo de aprendizaje supervisado donde el valor predicho es generado a través de una sucesión de longitud fija de pasos

$$x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_T \quad x_{n+1} = f_n(x_n, \theta_n)$$

donde las  $f_i$  son funciones predeterminadas que dependen de parámetros  $\theta_i$ . Los parámetros  $\theta_i$  también son conocidos como pesos. Cada paso del proceso se le conoce como capa. El valor

$x_T$  es el **valor predicho** por la red  $\{f_i(\cdot, \theta_i)\}$  para el valor de entrada  $x_0$ . Además, el valor  $x_n$  es el valor de entrada de la capa  $n$ , y el valor  $x_{n+1}$  su valor de salida. Comúnmente la  $i$ -ésima coordenada del paso  $k$

$$x_{k+1}^i = f_k^i(x_k, \theta_k)$$

es de la forma

$$f_k^i(x_k, \theta_k) = \sigma_k^i(W_k^i x_k + b_k^i) \quad \theta_k^i = (W_k^i, b_k^i), \quad \theta_k = (\theta_k^1, \dots, \theta_k^{n_k})$$

es decir, primero una transformación lineal, y luego se evalúa en una función  $\sigma_k^i$ , la función de activación. Las funciones de activación más usadas son

$$\begin{aligned} \text{ReLU}(x) &:= \max(x, 0) \\ \text{sigmoid}(x) &:= \frac{1}{1 + \exp\{-x\}} \\ \text{tanh}(x) &= \frac{1 - \exp\{-2x\}}{1 + \exp\{-2x\}} \\ \text{identidad} & \\ \text{sgn.} & \end{aligned}$$

Por **red residual** se entenderá una red neuronal donde, renombrando, se tiene

$$x_{k+1} = x_k + f_k(x_k, \theta_k).$$

Por **entrenamiento** se entenderá el proceso automatizado mediante el cual se obtienen valores para los parámetros de la red de modo que ésta prediga de forma eficiente.

La forma de medir eficiencia o idealidad de los parámetros, y eficiencia de la predicción, de una red neuronal, es a través de una **función objetivo**, la cual suele ser una suma de dos términos

$$J(\theta) = \Phi(x_T, y) + \sum_{n=0}^{T-1} L_n(\theta_n), \quad \theta = (\theta_0, \dots, \theta_{T-1}).$$

El término  $\Phi$  es la función de pérdida, y las funciones  $L_i$  se llaman regularizadores. Dichas funciones, pérdida y regularizadores, son funciones previamente escogidas. Habitualmente la función de pérdida es intrínseca al problema, y las de regularización al algoritmo de aprendizaje. Las funciones más usadas son

$$\begin{aligned} \Phi(x_T, y) &= |x_T - y|_1, |x_T - y|_2, 0-1 \text{ loss}(x_T, y) := \mathbb{1}_{\{y\}}(x_T) \\ L &= |\cdot|_1, |\cdot|_2. \end{aligned}$$

En la mayoría de los casos el término de regularización es para limitar que los parámetros  $\theta_i$  sólo tomen valores dentro un posible conjunto de valores  $\Theta$ .

Por otro lado, lo común es que las funciones involucradas en las capas de una red neuronal, la función de pérdida y las de regularización son funciones suaves, o diferenciables salvo en un conjunto determinado y conocido, y con propiedades que no presentan problemas numéricos para los algoritmos de optimización.



De esta manera y bajo una perspectiva simplista, un modelo de aprendizaje supervisado, como las redes en consideración, cuenta con una muestra  $\{(x^i, y^i)\}_{i=1}^N$ , una función objetivo,  $J(\theta)$ , la cual es diferenciable, y como propósito de entrenamiento es encontrar  $\theta$  que minimice

$$J(\theta) = \frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(x^i, y^i, \theta)$$

donde  $\tilde{\Phi}(x^i, y^i, \theta) = \Phi(x_T^i, y^i) + \sum_{n=0}^{T-1} L_n(\theta_n)$ . Para lograrlo propone buscar  $\theta$  tal que

$$\nabla_{\theta} J = 0.$$

Esto mediante un método iterativo, conocido como **descenso de gradiente**, en el que se hacen propuestas de  $\theta$  según el siguiente algoritmo

---

**Algorithm 1** Descenso de gradiente
 

---

Inicializa  $\theta^0$

**for**  $k = 0$  **to** # iteraciones **do**

$$\theta^{k+1} \leftarrow \theta^k - \eta \nabla_{\theta} \left( \frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(x_i, y_i, \theta^k) \right).$$


---

El valor  $\eta$  se conoce como razón de aprendizaje; puede ser un valor fijo, depender del valor actual de  $\theta$  y/o los anteriores. Según la variante del algoritmo, principalmente la forma de calcular  $\eta$ , el procedimiento se conoce como Descenso de gradiente con momento, la versión de Nesterov, AdaGrad, RMSProp, Adam, entre otros.

En la mayoría de casos de aprendizaje máquina, el tamaño de la muestra,  $N$ , es muy grande, implicando que calcular el gradiente con toda la muestra no es computacionalmente eficiente. En virtud de

$$\nabla_{\theta} \frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(x_i, y_i, \theta) = \frac{1}{N} \sum_{i=1}^N \nabla_{\theta} \tilde{\Phi}(x_i, y_i, \theta)$$

y suponiendo que la muestra es independiente e idénticamente distribuida (iid), entonces el promedio de la derecha se puede considerar como una aproximación a la esperanza  $\mathbb{E}_{\mu}(\nabla_{\theta} \tilde{\Phi}(x, y, \theta))$ , donde  $\mu$  es la medida de distribución de la variable  $(x, y)$ . Por otra parte, en situaciones prácticas es común poder aproximar una esperanza tomando un promedio sobre una muestra no tan grande, por ejemplo de tamaño 128 o tal vez menor. De esta manera, se toma como estrategia estimar el gradiente  $\nabla_{\theta} J$  con una submuestra de tamaño  $n$  fijo de la muestra de entrenamiento, es decir la actualización se hace

$$\theta^{k+1} \leftarrow \theta^k - \eta \frac{1}{n} \sum_{(x,y) \in \text{submuestra}} \nabla_{\theta} \tilde{\Phi}(x, y, \theta^k).$$

A este método se le conoce como **descenso de gradiente estocástico**. Dependiendo de la fuente, a la muestra original de entrenamiento se le llama **batch** y a la submuestra **mini-batch**, otros usan ambos nombres indistintamente para la submuestra.

Por otra parte, en redes neuronales el número de parámetros, dimensión de  $\theta = (\theta_0, \dots, \theta_{T-1})$ , suele ser muy grande, comúnmente mayor a un millón, lo cual plantea la necesidad de calcular el gradiente  $\nabla_{\theta} J$  de manera eficiente. **Backpropagation** es el nombre del algoritmo que calcula dichas derivadas, en forma ordenada, a través de la regla de la cadena, comenzando por la última capa y procediendo hacia atrás.

Note que la regla de la cadena indica

$$\frac{\partial J}{\partial \theta_k} = \sum_j \frac{\partial J}{\partial x_{k+1}^j} \frac{\partial f_k^j}{\partial \theta_k^i}$$

y una expresión análoga para las parciales  $\frac{\partial J}{\partial x_k^j}$ . Si inicia calculando  $\nabla J$  por la última capa, entonces para calcular  $\nabla_{\theta_k} J$ , sólo es necesario computar  $\frac{\partial f_k^j}{\partial \theta_k}$ , pues  $\nabla_{x_{k+1}} J$  ya se tendrá. Mas aún, al ser las funciones  $f_i$  fijas o predeterminadas, si se tiene una expresión simbólica de  $\frac{\partial f_k^j}{\partial \theta_k}$ , entonces sólo se tendrá que evaluar dicha expresión a lo largo del entrenamiento. Por último, recordando que las funciones comúnmente usadas en la construcción y evaluación de eficiencia de una red neuronal son elementos de una lista particular de funciones, y observando que para las funciones de dicha lista su derivada puede escribirse como una expresión sencilla en términos de los argumentos de la función o la función evaluada, resulta entonces que no es necesario siquiera crear una nueva expresión simbólica para cada derivada  $\frac{\partial f_k^j}{\partial \theta_k}$ ,  $\frac{\partial J}{\partial x_{k+1}^j}$ , sino sólo identificar los términos que ya se calcularon y se necesitan para suputar  $\partial_{\theta_k} J$ . Esto puede automatizarse desde la creación de la red neuronal. Muchas librerías explotan estas observaciones en su implementación de backpropagation, haciendo la implementación eficiente en el uso de memoria y respecto al tiempo de ejecución.

## 1.2. Preliminares matemáticos

### Definición 1.2.1.

1. Un espacio medible  $(\Omega, \mathcal{F})$  se dice Borel isomorfo a otro espacio medible  $(\Omega', \mathcal{F}')$  si existe una biyección  $f : \Omega \rightarrow \Omega'$  con  $f$  y  $f^{-1}$  medibles.
2. Un espacio polaco es un espacio topológico homeomorfo a un espacio métrico, separable y completo.
3. Dado un espacio topológico  $\Omega$ , su  $\sigma$ -álgebra de Borel se denotará como  $\mathcal{B}(\Omega)$ .

**Teorema 1.2.1.** Cualesquiera dos espacios polacos no contables son Borel isomorfos.

Para una prueba consultar corolario 2.6 y teorema 2.10 en <https://ncatlab.org/nlab/show/Polish+space>.

**Definición 1.2.2.** Una medida  $\mu : \mathcal{F} \rightarrow [0, \infty]$  se dice sin átomos si para todo  $A \in \mathcal{F}$  de medida positiva, existe  $B$  subconjunto propio de  $A$  con medida positiva.

**Ejemplo 1.2.1.** La medida de Lebesgue en  $\mathcal{B}([0, 1])$  no tiene átomos.

Sea  $X$  un espacio polaco no contable. Denote por  $\mathcal{M}(X)$  el espacio de medidas de probabilidad Borelianas en  $X$ , es decir, definidas en  $\mathcal{B}(X)$ , y por  $L^0(\Omega, X)$  el espacio de funciones medibles  $\xi : \Omega \rightarrow X$ .

**Propiedad 1.2.1. (Representación de medidas)** Sea  $(\Omega, \mathcal{F}, \mathbb{P})$ , un espacio de probabilidad con  $\Omega$  un espacio polaco no contable,  $\mathcal{F} = \mathcal{B}(\Omega)$ , y  $\mathbb{P}$  sin átomos, y sea  $X$  otro espacio polaco no contable. Entonces, la función

$$\begin{aligned} L^0(\Omega, X) &\rightarrow \mathcal{M}(X) \\ \xi &\rightarrow \mathbb{P}_\xi \end{aligned}$$

es suprayectiva, donde  $\mathbb{P}_\xi$  es la medida inducida por  $\xi$  en  $X$ , es decir  $\mathbb{P}_\xi(B) = \mathbb{P}(\xi \in B)$ .

*Demostración.* Sin pérdida de generalidad suponga que los espacios son  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . Sea  $\mu \in \mathcal{M}(\mathbb{R})$  y  $F, G$  las funciones de distribución de  $\mathbb{P}$  y  $\mu$  respectivamente. Como  $\mathbb{P}$  no tiene átomos,  $F$  es continua y  $F_*(\mathbb{P}) := \mathbb{P}_F$ , la medida en  $\mathbb{R}$  asociada a  $F$ , es la medida de Lebesgue,  $\lambda$ , en  $\mathcal{B}([0, 1])$ . Por otra parte sea  $G^{-1}$  la función de cuantiles de  $G$ , es decir

$$G^{-1}(t) := \inf\{x : G(x) \geq t\}$$

entonces  $G_*^{-1}(\lambda) = \mu$ . Por lo tanto  $(G^{-1} \circ F)_*(\mathbb{P}) = \mu$ .  $\square$

Dado un espacio medible  $(X, \mathcal{F})$  se dirá que una medida  $\mu$  está definida en  $X$ , sobrentendiendo que el dominio de  $\mu$  es  $\mathcal{F}$ . En el caso que  $X$  sea un espacio topológico, se supondrá que la  $\sigma$ -álgebra que se está considerando es la de Borel de  $X$ .

Sea  $(X, d)$  un espacio métrico, separable y completo, e.g.  $(\mathbb{R}^n, |\cdot|_2)$ , denote por  $\mathcal{P}_2(X)$  el espacio de medidas de probabilidad  $\mu$  en  $X$  que son cuadrado integrables, es decir que

$$\int_X d^2(x_0, x) d\mu(x) < \infty$$

para algún, y por lo tanto para cualquier,  $x_0 \in X$ . Defina la **distancia 2 de Wasserstein** o 2 de Monge-Kantorovich en  $\mathcal{P}_2(X)$  mediante

$$W_2(\mu, \hat{\mu}) = \inf_{\gamma \in \pi(\mu, \hat{\mu})} \left[ \int_{X^2} d(x, y)^2 d\gamma(x, y) \right]^{1/2}$$

donde

$$\pi(\mu, \hat{\mu}) = \{\gamma \in \mathcal{M}(X^2) : \gamma(A \times X) = \mu(A), \gamma(X \times B) = \hat{\mu}(B) \quad \forall A, B \in \mathcal{B}(X)\}.$$

**Propiedad 1.2.2.** La **distancia 2 de Wasserstein** es una distancia en  $\mathcal{P}_2(X)$ .

Para una demostración diríjase al lema 5.2 de [Cardaliaguet, 2013].

Sea  $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ , un espacio de probabilidad con  $\Omega$  polaco, y  $\mathbb{P}$  sin átomos, el cual fijaremos de ahora en adelante. Por la propiedad de representación de medidas

$$W_2(\mu, \hat{\mu}) = \inf\{|\xi - \zeta|_2 : \xi, \zeta \in L^2(\Omega, X), \mathbb{P}_\xi = \mu, \mathbb{P}_\zeta = \hat{\mu}\}.$$

Dadas  $\mu, \hat{\mu} \in \mathbb{P}_2(X)$  y  $\xi, \zeta \in L^2(\Omega, X)$  se usará la notación  $(\xi, \zeta) \in \Pi(\mu, \hat{\mu})$  para indicar  $\mathbb{P}_{(\xi, \zeta)} \in \Pi(\mu, \hat{\mu})$ .

Dada una función  $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$ , su levantamiento es la función  $\tilde{u} : L^2(\Omega, \mathbb{R}^n) \rightarrow \mathbb{R}$  definida por

$$\tilde{u}(\xi) = u(\mathbb{P}_\xi).$$

Una función  $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$  se dice diferenciable en  $\mu \in \mathcal{P}_2(\mathbb{R}^n)$  si existe  $\xi \in L^2(\Omega, \mathbb{R}^n)$  con  $\mathbb{P}_\xi = \mu$  y  $\tilde{u}$  es Fréchet diferenciable en  $\xi$ , es decir, si existe una transformación lineal continua  $d_\xi \tilde{u} : L^2(\Omega, \mathbb{R}^n) \rightarrow \mathbb{R}$  tal que

$$\lim_{h \rightarrow 0} \frac{|\tilde{u}(\xi + h) - \tilde{u}(\xi) - d_\xi \tilde{u}h|}{|h|} = 0$$

donde  $|\cdot|$  es la norma de  $L^2(\Omega, X)$ .

Una función  $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$  se dice de clase  $C^1$  si su levantamiento lo es.

Los siguientes resultados son utilizados en la demostración de existencia, unicidad y continuidad de soluciones de ecuaciones diferenciales ordinarias. Estos son centrales en la prueba de la propiedad 3.2.2 del capítulo 3, y debido a que en la referencia usada, [Bressan and Piccoli, 2007], se concentran en una suposición ligeramente distinta a la del presente trabajo, su demostración se incluye por completez.

**Proposición 1.2.1.** Sea  $0 < T < \infty$  fijo, y  $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  acotada por  $C$ , y Lipschitz continua uniformemente en  $\theta$ , es decir existe  $k$  tal que

$$|f(x, \theta) - f(y, \theta)| \leq k|x - y| \quad \forall x, y, \theta$$

entonces

i) La función  $|\cdot|_\dagger$  definida por

$$|x_t|_\dagger = \sup_{t \in [0, T]} e^{-2kt} |x_t|$$

es una norma en  $C([0, T], \mathbb{R}^m)$  equivalente a la norma del supremo (en el mismo espacio).

ii) Dado  $\theta_s \in L^\infty([0, T], \mathbb{R}^m)$  fijo, la transformación

$$x \mapsto \varphi(x, \theta) = a_0 + \int_0^t f(x_s, \theta_s) ds \quad t \in [0, T]$$

es contracción, con constante  $\kappa = \frac{1}{2}$ , en  $(C([0, T], \mathbb{R}^n), |\cdot|_\dagger)$ .

*Demostración.*

i) Para  $x \in C([0, T], \mathbb{R}^n)$

$$e^{-2KT} |x|_{L^\infty} \leq |x|_\dagger \leq |x|_{L^\infty}.$$

ii) Sean  $x, y \in C([0, T], \mathbb{R}^n)$ ,  $\theta \in L^\infty([0, T], \mathbb{R}^m)$ ,  $\delta = |x - y|_\dagger$ . Así  $|x_s - y_s| \leq e^{2ks} \delta$ , y la desigualdad

$$\begin{aligned} |\varphi(x, \theta) - \varphi(y, \theta)| &\leq \int_0^t |f(x_s, \theta_s) - f(y_s, \theta_s)| ds \\ &\leq \int_0^t k|x_s - y_s| ds \end{aligned}$$

implica

$$\begin{aligned} \sup_{t \in [0, T]} e^{-2kt} |\varphi(x, \theta)_t - \varphi(y, \theta)_t| &\leq \sup_{t \in [0, T]} \int_0^t e^{-2kt} k |x_s - y_s| ds \\ &\leq \sup_{t \in [0, T]} \int_0^t \delta k e^{2k(s-t)} ds \leq \frac{|x - y|_{\dagger}}{2}. \end{aligned}$$

□

**Proposición 1.2.2.** Sea  $X$  un espacio de Banach,  $\Lambda$  un conjunto, y sea

$$\varphi : X \times \Lambda \rightarrow X$$

tal que para algún  $\kappa < 1$

$$|\varphi(x, \lambda) - \varphi(y, \lambda)| \leq \kappa |x - y|$$

entonces para cada  $\lambda \in \Lambda$  existe un único punto fijo  $x_\lambda \in X$  tal que

$$x_\lambda = \varphi(x_\lambda, \lambda).$$

Además

$$|y - x_\lambda| \leq \frac{1}{1 - \kappa} |y - \varphi(y, \lambda)|.$$

*Demostración.* Sea  $y_0 \in X, \lambda \in \Lambda$ . Defina  $y_{n+1} = \varphi(y_n, \lambda), n \geq 0$ . La desigualdad

$$|y_{n+2} - y_{n+1}| = |\varphi(y_{n+1}, \lambda) - \varphi(y_n, \lambda)| \leq \kappa |y_{n+1} - y_n|$$

implica  $\{y_n\}$  es sucesión de Cauchy, y como  $X$  es Banach, entonces converge a algo,  $x_\lambda$ . Por continuidad de  $\varphi(\cdot, \lambda)$ , pues es contracción,  $x_\lambda = \lim_{n \rightarrow \infty} y_{n+1} = \varphi(\lim_{n \rightarrow \infty} y_n, \lambda)$ ,  $x_\lambda$  es punto fijo de  $\varphi(\cdot, \lambda)$ . También por ser  $\varphi(\cdot, \lambda)$  contracción, el punto fijo de  $\varphi(\cdot, \lambda)$  es único.

Por otra parte

$$|y - x_\lambda| \leq |y - \varphi(y, \lambda)| + |\varphi(y, \lambda) - \varphi(x_\lambda, \lambda)| \leq |y - \varphi(y, \lambda)| + \kappa |y - x_\lambda|$$

implicando

$$|y - x_\lambda| \leq \frac{1}{1 - \kappa} |y - \varphi(y, \lambda)|.$$

□



# Capítulo 2

## Problema de entrenamiento como problema de control

El objetivo de este capítulo es rescatar y detallar algunas de las demostraciones de los resultados enunciados en [E et al., 2018]. Éstos son sobre las propiedades matemáticas de la formulación del problema de entrenamiento de una red residual como problema de control óptimo. Dicho problema de control puede estudiarse desde dos perspectivas: Programación Dinámica y el Principio del Máximo de Pontryagin. El capítulo está constituido de la siguiente manera. En la primera sección se formula el problema de entrenamiento como problema de control óptimo, también se establece la notación empleada en el resto del trabajo. En la siguiente sección se establecen algunas propiedades del problema de control desde el punto de Programación Dinámica. En la última sección se establece y demuestra un Principio del Máximo de Pontryagin para el problema de control.

### 2.1. Formulación matemática

El problema de entrenamiento de una red residual de  $T$  capas, dada una muestra de entrenamiento  $\{(x_0^i, y_0^i)\}_{i=1}^N$ , consiste en encontrar  $\theta = (\theta_0, \dots, \theta_{T-1})$  que minimice

$$\frac{1}{N} \sum_{i=1}^N \left[ \Phi(x_T^i, y_0^i) + \sum_{k=0}^{T-1} L_k(\theta) \right]$$

sujeto a  $x_{k+1} = x_k + f_k(x_k, \theta_k)$ ,  $k = 0, \dots, T - 1$ .

Bajo la suposición que  $f_k$  es una función fija  $f$  en la que sólo cambian los parámetros  $\theta_k$ , la restricción

$$x_{k+1} = x_k + f(x_k, \theta_k)$$

puede interpretarse como una discretización de una ecuación diferencial donde los pesos  $\theta_k$  son parámetros de control, y un control  $\theta = \theta_t, t \in [0, T]$ , como una función  $\theta \in L^\infty([0, \infty], \Theta)$ <sup>1</sup>.

---

<sup>1</sup>Aquí y en el resto de la sección

$$L^\infty([0, T], \Theta) = \{\theta \in L^\infty([0, T], \mathbb{R}^m) : \theta_t \in \Theta \quad \forall t \in [0, T]\}.$$

Además, suponiendo  $L_i = L$ , es decir la misma función de regularización, la suma  $\sum_i L_i$  se interpretaría como una integral. Esto conlleva a estudiar el **problema empírico**

$$\inf_{\theta \in L^\infty([0, T], \Theta)} J_N(\theta) := \frac{1}{N} \sum_{i=1}^N \left[ \Phi(x_T^i, y_0^i) + \int_0^T L(x_t^i, \theta_t) dt \right] \quad (2.1)$$

sujeto a  $\dot{x}_t^i = f(x_t^i, \theta_t) \quad t \in [0, T], i = 1, \dots, N$

donde  $(x_0^i, y_0^i)$  es una muestra iid con medida de distribución  $\mu_0$ ; y también el **problema teórico**

$$\inf_{\theta \in L^\infty([0, T], \Theta)} J(\theta) := \mathbb{E}_{\mu_0} \left[ \Phi(x_T, y_0) + \int_0^T L(x_t, \theta_t) dt \right] \quad (2.2)$$

sujeto a  $\dot{x}_t = f(x_t, \theta_t) \quad t \in [0, T]$ .

Cabe resaltar que el control óptimo que se busca es conocido desde el comienzo, es decir no tiene retroalimentación del estado  $x_t$ .

Note que en general la solución del problema empírico dependerá de la muestra y potencialmente será una variable aleatoria.

La siguiente es la notación usada a lo largo del trabajo, a menos que se indique lo contrario:  $x \in \mathbb{R}^d, y \in \mathbb{R}^l, w = (x, y) \in \mathbb{R}^{d+l}, \Theta \subset \mathbb{R}^m$  es el espacio de posibles valores de  $\theta_t, T > 0$  es fijo. Las funciones  $f, \Phi, L$  tienen dominios respectivos

$$\begin{aligned} f &: \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}^d \\ \Phi &: \mathbb{R}^d \times \mathbb{R}^l \rightarrow \mathbb{R} \\ L &: \mathbb{R}^d \times \Theta \rightarrow \mathbb{R} \end{aligned}$$

$$\begin{aligned} \bar{f}(w, \theta) &:= (f(x, \theta), 0) \in \mathbb{R}^{d+l} \\ \bar{L}(w, \theta) &:= L(x, \theta) \\ \bar{\Phi}(w) &:= \Phi(x, y). \end{aligned}$$

El producto interno en un espacio Euclidiano es  $\langle \cdot, \cdot \rangle$ . Además, se fijará un espacio de probabilidad  $(\Omega, A, \mathbb{P})$  tal que toda medida cuadrado integrable  $\mu \in P_2(\mathbb{R}^{d+l})$  se pueda inducir por una variable aleatoria en  $\Omega$ , i.e.  $\mu = \mathbb{P}_X$  para alguna  $X \in L^0(\Omega, \mathbb{R}^{d+l})$ , por ejemplo  $\Omega$  podría ser un espacio polaco,  $A = \mathcal{B}(\Omega)$  su  $\sigma$  álgebra de Borel y  $\mathbb{P}$  una medida de probabilidad sin átomos.

Además,  $L^2(\Omega, \mathbb{R}^{d+l})$  denotará el espacio de Hilbert de variables aleatorias  $\mathbb{R}^{d+l}$ -valuadas definidas en  $\Omega$  que son cuadrado integrables. El producto interno es  $\langle \xi, \hat{\xi} \rangle_{L^2} = \mathbb{E}(\xi(\omega), \hat{\xi}(\omega))$ .

Dada una función medible  $\psi : \mathbb{R}^{d+l} \rightarrow \mathbb{R}^q$  que es cuadrado integrable con respecto a  $\mu \in P_2(\mathbb{R}^{d+l})$ , la notación  $\langle \psi, \mu \rangle$  representará

$$\langle \psi, \mu \rangle = \int_{\mathbb{R}^{d+l}} \psi(x) d\mu(x).$$

Dado  $\xi \in L^2(\Omega, \mathbb{R}^{d+l})$ , control  $\theta \in L^\infty([0, T], \Theta)$ , y  $0 \leq t \leq \hat{t} \leq T$ ,  $W_t^{t, \xi, \theta}$  representará el sistema dinámico en  $L^2(\Omega, \mathbb{R}^{d+l})$  dado por

$$W_{\hat{t}}^{t, \xi, \theta} = \xi + \int_t^{\hat{t}} \bar{f}(W_s^{t, \xi, \theta}, \theta_s) ds.$$



Note que para  $\mu \in P_2(\mathbb{R}^{d+l})$  y  $\xi \in L^2(\Omega, \mathbb{R}^{d+l})$  con  $\mathbb{P}_\xi = \mu$ , la ley de  $W_s^{t,\xi,\theta}$  depende de  $\mu$  pero no de la variable  $\xi$  si  $\bar{f}$  es acotada y uniformemente Lipschitz en  $w$  (considere el flujo  $h(s, t, w_0)$  de la ecuación diferencial  $\frac{dw}{ds} = \bar{f}(w(s), \theta_s)$ ,  $w(t) = w_0$ . Éste existe por las condiciones en  $f$ . Por cambio de medida  $\mathbb{P}_{W_s^{t,\xi,\theta}} = h(s, t, \xi)\mu$ ). Así, es justificado escribir la ley de  $W_s^{t,\xi,\theta}$  como

$$\mathbb{P}_s^{t,\mu,\theta} := \mathbb{P}_{W_s^{t,\xi,\theta}}.$$

Más aún, la propiedad de flujo o semigrupo de la solución de la ecuación diferencial implica

$$\mathbb{P}_t^{t,\mu,\theta} = \mathbb{P}_t^{s,\mathbb{P}_s^{t,\mu,\theta},\theta}.$$

Finalmente las letras  $C, k$  representarán constantes que puede cambiar de línea a línea. Usualmente,  $C$  se usará para una cota, con un subíndice que indique el objeto que acota, y  $k$  para la constante en funciones Lipschitz. Cuando sólo se use la letra  $C, k$ , ésta puede cambiar de renglón a renglón.

## 2.2. Principio de Programación Dinámica

En esta sección se estudiarán algunas propiedades del problema teórico desde el punto de vista de Programación Dinámica. El objetivo último sería demostrar que la función de valor, definida más adelante, es la única solución de viscosidad de la ecuación de Hamilton-Jacobi-Bellman. A pesar que este resultado no se recuperó, sí fue posible demostrar algunos otros. Los que resultados que se detallan son: un resultado sobre continuidad Lipschitz de la función objetivo  $J$ , el Principio de Programación Dinámica para la función de valor  $v$ , y una proposición referente a la ecuación de Hamilton-Jacobi-Bellman.

A lo largo de esta subsección se supondrá  $\bar{f}, \bar{L}, \bar{\Phi}$  son acotadas por  $C_f, C_L, C_\Phi$ , y Lipschitz continuas con constantes  $K_f, K_L, K_\Phi$ . Además  $\bar{f}, \bar{L}$  son uniformemente Lipschitz continuas, es decir la constante de Lipschitz continuidad no depende de  $\theta$ . Por último, el valor  $n$  es igual a  $d + l$ .

Dada  $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ ,  $\theta \in L^\infty([0, T], \Theta)$ ,  $0 \leq t \leq T$ , defina

$$J(t, \mu, \theta) = \mathbb{E}_{(x_t, y) \sim \mu} \left[ \Phi(x_T, y) + \int_t^T L(x_s, \theta_s) ds \right].$$

Con la notación indicada en preliminares

$$\begin{aligned} \mathbb{E}_{(x_t, y) \sim \mu} \Phi(x_T, y) &= \langle \Phi, \mathbb{P}_T^{t,\mu,\theta} \rangle \\ \mathbb{E}_\mu \int_t^T L(x_s, \theta_s) ds &= \int_t^T \mathbb{E} L(x_s, \theta_s) ds \\ &= \int_t^T \langle \bar{L}(\cdot, \theta_s), \mathbb{P}_s^{t,\mu,\theta} \rangle ds. \end{aligned}$$

**Definición 2.2.1.** La función de valor  $v^*(t, \mu)$ ,  $t \in [0, T]$ ,  $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ , se define como

$$v^*(t, \mu) := \inf_{\theta \in L^\infty([0, T], \Theta)} J(t, \mu, \theta).$$

Note que en la definición de función de valor sólo interesa el comportamiento de los controles en el intervalo  $[t, T]$ . Fijando un punto  $w \in \Theta$ , un control  $\theta \in L^\infty([t, T], \Theta)$  naturalmente se puede extender a un control en  $[0, T]$ , definiéndolo como  $\theta_s = w$  para  $0 \leq s < t$ . De ahí que en la definición el ínfimo se tome sobre controles en  $[0, T]$ .

**Propiedad 2.2.1.** La función  $(t, \mu) \mapsto J(t, \mu, \theta)$  es Lipschitz continua en  $[0, T] \times \mathcal{P}_2(\mathbb{R}^n)$ , y la constante de Lipschitz no depende de  $\theta$ .

*Demostración.* Observe

$$\begin{aligned} |\langle \bar{L}(\cdot, \theta), \mu \rangle - \langle \bar{L}(\cdot, \theta), \hat{\mu} \rangle| &= |\mathbb{E}[L(X, \theta) - L(Y, \theta)]| \quad (X, Y) \in \Pi(\mu, \hat{\mu}) \\ &\leq K_L \mathbb{E}|X - Y|_{\mathbb{R}^n} \quad \text{Lipschitz} \\ &\leq K_L |X - Y|_{L^2} \quad \text{concavidad de } \sqrt{z}. \end{aligned}$$

Esto ocurre para todo  $(X, Y) \in \Pi(\mu, \hat{\mu})$ ; así tomando el ínfimo se obtiene

$$|\langle \bar{L}(\cdot, \theta), \mu \rangle - \langle \bar{L}(\cdot, \theta), \hat{\mu} \rangle| \leq K_L W_2(\mu, \hat{\mu}).$$

Por un argumento equivalente

$$|\langle \bar{\Phi}(\cdot), \mu \rangle - \langle \bar{\Phi}(\cdot), \hat{\mu} \rangle| \leq K_\Phi W_2(\mu, \hat{\mu}).$$

Dado  $\theta \in L^\infty([0, T], \Theta)$  defina  $h(t, t_0, w_{t_0})$  la solución de

$$\begin{aligned} \frac{dw^\theta}{dt} &= \bar{f}(w^\theta(t), \theta_t), \quad t_0 \leq t \leq T \\ w^\theta(t_0) &= w_{t_0}. \end{aligned}$$

Entonces

$$\begin{aligned} |h(t, t_0, w) - h(t, t_0, \hat{w})| &\leq |w - \hat{w}| + \int_0^t |\bar{f}(h(s, t_0, w), \theta_s) - \bar{f}(h(s, t_0, \hat{w}), \theta_s)| ds \\ &\leq |w - \hat{w}| + k_f \int_{t_0}^t |h(s, t_0, w) - h(s, t_0, \hat{w})| \\ &\leq |w - \hat{w}| \exp\{k_f T\} \quad \text{desigualdad de Gronwall} \end{aligned}$$

y para  $t_0 \leq t \leq \hat{t} \leq T$

$$\begin{aligned} |h(t, t_0, w) - h(\hat{t}, t_0, w)| &\leq \int_t^{\hat{t}} |\bar{f}(h(s, t_0, w), \theta_s)| ds \\ &\leq C_f |t - \hat{t}|. \end{aligned}$$

Así,

$$\begin{aligned} W_2(\mathbb{P}_s^{t, \mu, \theta}, \mathbb{P}_s^{t, \hat{\mu}, \theta}) &= \inf\{|X - Y|_{L^2} : (X, Y) \in \Pi(\mathbb{P}_s^{t, \mu, \theta}, \mathbb{P}_s^{t, \hat{\mu}, \theta})\} \\ &\leq \inf\{|h(s, t, X) - h(s, t, Y)|_{L^2} : (X, Y) \in \Pi(\mu, \hat{\mu})\} \\ &\leq K_h W_2(\mu, \hat{\mu}). \end{aligned}$$

Análogamente

$$\begin{aligned} W_2(\mathbb{P}_s^{t,\mu,\theta}, \mu) &\leq \inf\{|h(s, t, X) - X|_{L^2} : \mathbb{P}_X = \mu\} \\ &\leq C_f |s - t|. \end{aligned}$$

De este modo, para  $0 \leq t \leq \hat{t} \leq s \leq T$

$$\begin{aligned} W_2(\mathbb{P}_s^{t,\mu,\theta}, \mathbb{P}_s^{\hat{t},\hat{\mu},\theta}) &= W_2(\mathbb{P}_s^{\hat{t},\mathbb{P}_t^{t,\mu,\theta},\theta}, \mathbb{P}_s^{\hat{t},\hat{\mu},\theta}) \\ &\leq KW_2(\mathbb{P}_t^{t,\mu,\theta}, \hat{\mu}) \\ &\leq K(W_2(\mathbb{P}_t^{t,\mu,\theta}, \mathbb{P}_t^{\hat{t},\hat{\mu},\theta}) + W_2(\mathbb{P}_t^{\hat{t},\hat{\mu},\theta}, \hat{\mu})) \\ &\leq K(W_2(\mu, \hat{\mu}) + |t - \hat{t}|). \end{aligned}$$

Por lo tanto, para  $0 \leq t \leq \hat{t} \leq T, \mu, \hat{\mu} \in \mathcal{P}_2(\mathbb{R}^n), \theta \in L^\infty([0, T], \Theta)$  se cumple

$$\begin{aligned} |J(t, \mu, \theta) - J(\hat{t}, \hat{\mu}, \theta)| &\leq |\langle \bar{\Phi}, \mathbb{P}_T^{t,\mu,\theta} \rangle - \langle \bar{\Phi}, \mathbb{P}_T^{\hat{t},\hat{\mu},\theta} \rangle| + \int_t^{\hat{t}} |\langle \bar{L}(\cdot, \theta_s, \mathbb{P}_s^{t,\mu,\theta}) \rangle| \\ &\quad + \int_{\hat{t}}^T |\langle \bar{L}(\cdot, \theta_s, \mathbb{P}_s^{t,\mu,\theta}) \rangle - \langle \bar{L}(\cdot, \theta_s, \mathbb{P}_s^{\hat{t},\hat{\mu},\theta}) \rangle| ds \\ &\leq (K_\Phi K)[W_2(\mu, \hat{\mu}) + |t - \hat{t}|] + C_f |\hat{t} - t| \\ &\quad + TK_L(W_2(\mu, \hat{\mu}) + |t - \hat{t}|) \\ &\leq K(W_2(\mu, \hat{\mu}) + |t - \hat{t}|). \end{aligned}$$

□

La importancia de la proposición anterior es establecer una fuerte propiedad de continuidad en  $J(t, \mu, \theta)$ . Se espera que la función de valor herede la continuidad uniforme. Esto último es el primer paso para la verificación de que la función de valor es solución de viscosidad de la ecuación de Hamilton Jacobi Bellman. Desafortunadamente este paso no se logró recuperar.

Por otra parte, la función de valor tiene una propiedad importante conocida como **Principio de Programación Dinámica**. El enunciado matemático de éste es la siguiente proposición.

**Proposición 2.2.1** (Principio de Programación Dinámica). Para  $0 \leq t \leq \hat{t} \leq T, \mu \in \mathcal{P}_2(\mathbb{R}^n)$  se cumple

$$v^*(t, \mu) = \inf_{\theta \in L^\infty([0, T], \Theta)} \int_t^{\hat{t}} \langle \bar{L}(\cdot, \theta_s), \mathbb{P}_s^{t,\mu,\theta} \rangle ds + v^*(\hat{t}, \mathbb{P}_{\hat{t}}^{t,\mu,\theta}).$$

*Demostración.* Dado  $0 \leq t \leq \hat{t} \leq T, \mu, \theta^1 \in L^\infty([0, T], \Theta), \epsilon > 0$ , considere control  $\theta^2$  tal que

$$v^*(\hat{t}, \mathbb{P}_{\hat{t}}^{t,\mu,\theta^1}) + \epsilon \geq \langle \bar{\Phi}, \mathbb{P}_T^{\hat{t},\mathbb{P}_{\hat{t}}^{t,\mu,\theta^1},\theta^2} \rangle + \int_{\hat{t}}^T \langle \bar{L}(\cdot, \theta_s^2), \mathbb{P}_s^{\hat{t},\mathbb{P}_{\hat{t}}^{t,\mu,\theta^1},\theta^2} \rangle ds.$$

Sea  $\hat{\theta}$  definido por

$$\hat{\theta}_s = 1_{\{s < \hat{t}\}} \theta_s^1 + 1_{\{s \geq \hat{t}\}} \theta_s^2$$

entonces

$$\begin{aligned}
 v^*(t, \mu) &\leq \int_t^T \langle \bar{L}(\cdot, \hat{\theta}_s), \mathbb{P}_s^{t, \mu, \hat{\theta}} \rangle ds + \langle \bar{\Phi}, \mathbb{P}_T^{t, \mu, \hat{\theta}} \rangle \\
 &= \int_t^{\hat{t}} \langle \bar{L}(\cdot, \theta_s^1), \mathbb{P}_s^{t, \mu, \theta^1} \rangle ds + \int_{\hat{t}}^T \langle \bar{L}(\cdot, \theta_s^2), \mathbb{P}_s^{\hat{t}, \mathbb{P}_{\hat{t}}^{t, \mu, \theta^1}, \theta^2} \rangle ds + \langle \bar{\Phi}, \mathbb{P}_T^{\hat{t}, \mathbb{P}_{\hat{t}}^{t, \mu, \theta^1}, \theta^2} \rangle \\
 &\leq \int_t^{\hat{t}} \langle \bar{L}(\cdot, \theta_s^1), \mathbb{P}_s^{t, \mu, \theta^1} \rangle ds + v^*(\hat{t}, \mathbb{P}_{\hat{t}}^{t, \mu, \theta^1}) + \epsilon.
 \end{aligned}$$

Tomando el ínfimo

$$v^*(t, \mu) \leq \inf_{\theta} \left[ \int_t^{\hat{t}} \langle \bar{L}(\cdot, \theta_s), \mathbb{P}_s^{t, \mu, \theta} \rangle ds + v^*(\hat{t}, \mathbb{P}_{\hat{t}}^{t, \mu, \theta}, \theta) \right].$$

Por otra parte, sea  $\epsilon > 0$ ,  $\theta^3 \in L^\infty([0, T], \Theta)$  tal que

$$v^*(t, \mu) + \epsilon \geq \langle \bar{\Phi}, \mathbb{P}_T^{t, \mu, \theta^3} \rangle + \int_t^T \langle \bar{L}(\cdot, \theta_s^3), \mathbb{P}_s^{t, \mu, \theta^3} \rangle ds$$

entonces

$$\begin{aligned}
 v^*(t, \mu) + \epsilon &\geq \langle \bar{\Phi}, \mathbb{P}_T^{\hat{t}, \mathbb{P}_{\hat{t}}^{t, \mu, \theta^3}, \theta^3} \rangle + \int_{\hat{t}}^T \langle \bar{L}(\cdot, \theta_s^3), \mathbb{P}_s^{\hat{t}, \mathbb{P}_{\hat{t}}^{t, \mu, \theta^3}, \theta^3} \rangle ds \\
 &\quad + \int_t^{\hat{t}} \langle \bar{L}(\cdot, \theta_s^3), \mathbb{P}_s^{t, \mu, \theta} \rangle ds \\
 &\geq v^*(\hat{t}, \mathbb{P}_{\hat{t}}^{t, \mu, \theta^3}) + \int_t^{\hat{t}} \langle \bar{L}(\cdot, \theta_s^3), \mathbb{P}_s^{t, \mu, \theta^3} \rangle ds \\
 &\geq \inf_{\theta} \left[ \int_t^{\hat{t}} \langle \bar{L}(\cdot, \theta_s), \mathbb{P}_s^{t, \mu, \theta} \rangle ds + v^*(\hat{t}, \mathbb{P}_{\hat{t}}^{t, \mu, \theta}, \theta) \right].
 \end{aligned}$$

□

Coloquialmente este principio indica que para alcanzar optimalidad desde el comienzo se puede buscar un punto intermedio en el tiempo, encontrar el estado óptimo en ese momento, y luego ir de forma eficiente del inicio a ese estado en la cantidad de tiempo correspondiente. Así, desde un punto de vista computacional, en el que se discretiza el tiempo, de manera ingenua se podría intentar encontrar el control óptimo de la siguiente forma: tome el penúltimo tiempo,  $t_{(n-1)}$  por llamarlo así, luego calcule la posición óptima para ese tiempo, después encuentre una manera de ir del estado inicial a esa posición en tiempo  $t_{(n-1)}$ , y así sucesivamente renombrando  $T = t_{(n-1)}$ . Desafortunadamente dicha estrategia no es computacionalmente posible, ya que implica el análisis de cada posible estado en cada momento del tiempo.

Por otra parte, en forma 'infinitesimal' el Principio de Programación Dinámica conlleva a que la función de valor satisface una ecuación diferencial parcial. Para establecer dicha ecuación se necesita un resultado sobre la estructura de las funciones suaves en  $\mathcal{P}_2(\mathbb{R}^n)$  que permitirá definir y tomar las derivadas correspondientes de funciones definidas en dicho espacio.

Dada una función  $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$  su levantamiento o extensión  $\tilde{u}$  a  $L^2(\Omega, \mathbb{R}^n)$  se define como

$$\tilde{u}(\xi) = u(\mathbb{P}_\xi).$$

Una función  $u$  en  $\mathcal{P}_2(\mathbb{R}^n)$  se dirá que es  $C^1$  o que  $u \in C^1(\mathcal{P}_2(\mathbb{R}^n))$  si su levantamiento a  $L^2(\Omega, \mathbb{R}^n)$  lo es.

**Teorema 2.2.1** (Estructura de gradiente). Sea  $u : \mathcal{P}_2(\mathbb{R}^n)$  de clase  $C^1$ ,  $\tilde{u} : L^2(\Omega, \mathbb{R}^n) \rightarrow \mathbb{R}$  su levantamiento, y  $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ , entonces existe una función  $\partial_\mu u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$  tal que

$$\nabla_\xi \tilde{u} = \partial_\mu u(\xi) \quad \text{c.s.} \quad \forall \xi \in L^2(\Omega, \mathbb{R}^n) \text{ con } \mathbb{P}_\xi = \mu.$$

Para la demostración de este resultado diríjase a la proposición 5.25 de [Carmona and Delarue, 2018].

Con la notación anterior, dado  $\theta \in L^\infty([0, T], \Theta)$ ,  $v \in C^1([0, T] \times \mathcal{P}_2(\mathbb{R}^n))$  y  $W_t^{t_0, \mu, \theta}$  definido por

$$W_t^{t_0, \mu, \theta} = W^{t_0} + \int_{t_0}^t \bar{f}(W_s^{t_0, \mu, \theta}, \theta_s) ds \quad \mathbb{P}_{W^{t_0}} = \mu$$

se cumple

$$\begin{aligned} \frac{d}{dt} v(t, \mathbb{P}_t^{t_0, \mu, \theta}) &= \frac{d}{dt} \tilde{v}(t, W_t^{t_0, \mu, \theta}) \\ &= \partial_t \tilde{v}(t, W_t^{t_0, \mu, \theta}) + \langle \partial_\mu \tilde{v}(t, W_t^{t_0, \mu, \theta}), \bar{f}(W_t^{t_0, \mu, \theta}) \rangle_{L^2} \\ &= \partial_t v(t, W_t^{t_0, \mu, \theta}) + \mathbb{E} \langle \partial_\mu \tilde{v}(t, W_t^{t_0, \mu, \theta}), \bar{f}(W_t^{t_0, \mu, \theta}) \rangle_{\mathbb{R}^n} \\ &= \partial_t v(t, W_t^{t_0, \mu, \theta}) + \mathbb{E} \langle \partial_\mu v(t, \mathbb{P}_t^{t_0, \mu, \theta})(W_t^{t_0, \mu, \theta}), \bar{f}(W_t^{t_0, \mu, \theta}) \rangle_{\mathbb{R}^n} \\ &= \partial_t v(t, W_t^{t_0, \mu, \theta}) + \langle \partial_\mu v(t, \mathbb{P}_t^{t_0, \mu, \theta}) \cdot \bar{f}(\cdot), \mathbb{P}_t^{t_0, \mu, \theta} \rangle. \end{aligned}$$

Por otra parte, de la igualdad dada por el Principio de Programación Dinámica

$$0 = \inf_{\theta \in L^\infty([0, T], \Theta)} \left[ v^*(t + \delta t, \mathbb{P}_{t+\delta t}^{t, \mu, \theta}) - v^*(t, \mu) + \int_t^{t+\delta t} \langle \bar{L}(\cdot, \theta_s), \mathbb{P}_s^{t, \mu, \theta} \rangle ds \right]$$

se imagina que al tomar el límite  $\delta t \downarrow 0$  se cumpla

$$\frac{\partial v}{\partial t} + \inf_{\theta \in L^\infty([0, T], \Theta)} \langle \partial_\mu v(t, \mu)(\cdot) \cdot \bar{f}(\cdot, \theta_t) + \bar{L}(\cdot, \theta_t), \mu \rangle = 0 \quad \text{en } [0, T] \times \mathcal{P}_2(\mathbb{R}^n) \quad (\text{HJB})$$

$$v(T, \mu) = \langle \bar{\Phi}, \mu \rangle \quad \text{en } \mu \in \mathcal{P}_2(\mathbb{R}^n).$$

Dicha ecuación (el par) se referirá como ecuación de Hamilton-Jacobi-Bellman o HJB por comodidad. Note que en cada punto  $t$  se toma el ínfimo sobre los controles  $\theta$ , sin embargo, además de la facilidad de notación permitida por la existencia de  $\partial_\mu v$  definida por el teorema (2.2.1), la existencia de  $\partial_\mu v$  indica que el ínfimo en HJB es puntual, es decir, se podría tomar el ínfimo sobre  $w \in \Theta$ , colocando  $w$  en las respectivas funciones.

La importancia de la ecuación HJB es parcialmente explicada por la siguiente proposición.

**Proposición 2.2.2.** Sea  $v \in C^1([0, T] \times \mathcal{P}_2(\mathbb{R}^n))$  una solución de (HJB), y suponga que para todo  $(t, \mu)$  existe control  $\theta^\dagger(t, \mu)$  que realiza el ínfimo de HJB, entonces  $v = v^*$  y

$$\theta = \theta^\dagger(t, \mathbb{P}_t^{0, \mu, \theta})$$

es control óptimo para el problema de entrenamiento.

*Demostración.* Para cualquier control  $\theta \in L^\infty([0, T], \Theta)$  se cumple

$$\begin{aligned} v(t, \mu) &= v(T, \mathbb{P}_T^{t, \mu, \theta}) - \int_t^T \partial_t v(t, \mathbb{P}_s^{t, \mu, \theta}) + \langle \partial_\mu v(t, \mathbb{P}_s^{t, \mu, \theta}) \cdot \bar{f}(\cdot), \mathbb{P}_s^{t, \mu, \theta} \rangle ds \\ &\leq \langle \bar{\Phi}, \mathbb{P}_T^{t, \mu, \theta} \rangle + \int_t^T \langle \bar{L}(\cdot, \theta_s), \mathbb{P}_s^{t, \mu, \theta} \rangle ds \\ &= J(t, \mu, \theta). \end{aligned}$$

Tomando el ínfimo sobre  $\theta$

$$v(t, \mu) \leq v^*(t, \mu).$$

Usando  $\theta = \theta^\dagger(t, \mathbb{P}_s^{t, \mu, \theta})$  se tiene la igualdad  $v(t, \mu) = J(t, \mu, \theta)$ , implicando

$$v(t, \mu) \geq \inf_{\theta} J(t, \mu, \theta) = v^*(t, \mu).$$

Por lo tanto  $v = v^*$ . □

Esta proposición indica que, suponiendo las condiciones de suavidad y existencia de controles óptimos, una solución de la ecuación de Hamilton Jacobi Bellman da el valor óptimo para la función objetivo  $J$  comenzando en un tiempo  $t$ . Además proporciona una manera teórica de hallar el control óptimo  $\theta$  que permite alcanzar el óptimo. Éste último, aunque no depende del estado  $x_t^\theta$ , es identificado por la evolución de la medida de distribución del par  $(x_t, y)$ , de modo que para hallar dicho control aún se necesita cierta información que no se tiene por completo desde el inicio, es decir al tiempo  $t = 0$ .

Como se mencionó al principio de la sección, el objetivo de estudiar el problema teórico desde la perspectiva de Programación Dinámica es demostrar que, aún sin la suposición de suavidad de la proposición anterior, la función de valor es en algún sentido la única solución de la ecuación de Hamilton-Jacobi-Bellman. En turno, y en conjunto con el Principio de Programación Dinámica, esto se puede interpretar como condiciones de suficiencia para que una función realice el mínimo de la función objetivo  $J$ . Sin embargo, en términos prácticos esto es de limitada utilidad. Aún suponiendo la suavidad suficiente, la ecuación HJB invita a resolver una ecuación diferencial parcial en un espacio de dimensión infinita para poder encontrar  $v$ . Al mismo tiempo, el Principio de Programación Dinámica no ofrece una metodología práctica realizable para hallar el control óptimo  $\theta$ , suponiendo que existe, ni tampoco una aproximación. De esta manera, para el interés de redes neuronales, es necesario recurrir a otra perspectiva.

## 2.3. Principio del Máximo de Pontryagin

En control óptimo determinista existe un resultado llamado **Principio del Máximo de Pontryagin** o PMP, el cual establece condiciones necesarias para que un control  $\theta$  sea óptimo. Éste está basado en ideas de Cálculo de variaciones, ver [Fleming and Rishel, 2012]. En éste, el problema general consiste en buscar una trayectoria, sujeta a unas restricciones en los extremos, que optimice un funcional  $\Phi$ . Para hallar condiciones de optimalidad una parte de la suposición que la trayectoria óptima existe, llámeme  $u_t^*$ , luego considera una familia  $u_t^\epsilon$ , con  $\epsilon$  en una vecindad del  $0 \in \mathbb{R}$ , de trayectorias con  $u_t^* = u_t^0$ . Tomando la derivada de  $\Phi(u_t^\epsilon)$  en dirección de  $\epsilon$  y evaluando en 0, es posible establecer un sistema de ecuaciones diferenciales, en dimensión finita, que la trayectoria óptima debe cumplir. En el Principio del Máximo de Pontryagin se parte de un razonamiento parecido, pero en vez de llegar a una ecuación diferencial se llega a una condición de maximalidad para cada tiempo  $t$ .

Por otra parte, el PMP mismo sugiere una manera de buscar un control que satisfaga dichas condiciones de necesidad. Así, la intención de esta sección es establecer un resultado análogo al PMP para el problema teórico. Para tal fin, a lo largo de la sección se supondrá

(A1) Las funciones  $f, L$  son acotadas;  $f, L$  son continuas en  $\theta$ , y  $f, L, \Phi$  son continuamente diferenciables con respecto a  $x$ .

(A2) La medida de distribución  $\mu_0$  tiene soporte acotado en  $\mathbb{R}^{d+l}$ , es decir, existe  $M > 0$  tal que

$$\mu_0(\{(x, y) \in \mathbb{R}^{d+l} : |x| + |y| \leq M\}) = 1.$$

**Teorema 2.3.1** (Principio del Máximo de Pontryagin). Suponga que se tienen las condiciones (A1), (A2) y que existe un control óptimo  $\theta^*$  para el problema teórico (2.2), es decir  $J(\theta^*)$  realiza el ínfimo. Entonces los procesos  $x^*, p^*$  definidos por

$$\dot{x}_t^* = f(x_t^*, \theta_t^*) \quad x_0^* = x_0 \quad (2.3)$$

$$\dot{p}_t^* = -\nabla_x H(x_t^*, p_t^*, \theta_t^*) \quad p_T = -\nabla_x \Phi(x_T^*, y_0) \quad (2.4)$$

donde  $(x_0, y_0) \sim \mu_0$ , satisfacen

$$\mathbb{E}_{\mu_0} H(x_t^*, p_t^*, \theta_t^*) \geq \mathbb{E}_{\mu_0} H(x_t^*, p_t^*, w) \quad \forall w \in \Theta, \text{ c.s. en } t \in [0, T]$$

donde el Hamiltoniano  $H : \mathbb{R}^d \times \mathbb{R}^d \times \Theta$  es definido por

$$H(x, p, w) = p \cdot f(x, w) - L(x, w).$$

*Demostración.* Introduzca la variable  $x_t^0 = \int_0^t L(x_t, \theta_t) dt$ . De este modo, renombrando

$$x \leftarrow (x^0, x) \quad f \leftarrow (L, f) \quad \Phi(x_T, y_0) \leftarrow \Phi(x_T, y_0) + x_T^0$$

podemos suponer sin pérdida de generalidad  $L \equiv 0$ , y retener las suposiciones (A1) y (A2).

Sea  $\tau \in (0, T]$  punto de Lebesgue<sup>2</sup> de  $t \rightarrow f(x_t^*, \theta_t^*)$ . Dado  $\epsilon \in (0, \tau]$ ,  $w \in \Theta$  defina

$$\theta_t^{\tau, \epsilon} = \begin{cases} w & t \in [\tau - \epsilon, \tau] \\ \theta_t^* & \text{otro caso} \end{cases}$$

$$x_t^{\tau, \epsilon} = x_0 + \int_0^t f(x_s^{\tau, \epsilon}, \theta_s^{\tau, \epsilon}) ds.$$

Note

$$x_t^{\tau, \epsilon} = x_t^* \quad t \leq \tau - \epsilon.$$

Así, para  $t = \tau$  se tiene

$$\lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} (x_\tau^{\tau, \epsilon} - x_\tau^*) = \frac{1}{\epsilon} f(x_\tau^{\tau, \epsilon}, w) - f(x_\tau^*, \theta_\tau^*).$$

A saber,

$$\frac{1}{\epsilon} (x_\tau^{\tau, \epsilon} - x_\tau^*) = \frac{1}{\epsilon} \int_{\tau - \epsilon}^{\tau} f(x_s^{\tau, \epsilon}, w) - f(x_s^*, \theta_s^*) ds.$$

Dado que  $\tau$  es punto de Lebesgue de  $t \mapsto f(x_t^*, \theta_t^*)$ , entonces

$$\lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \int_{\tau - \epsilon}^{\tau} f(x_s^*, \theta_s^*) ds = f(x_\tau^*, \theta_\tau^*).$$

Por otra parte, para  $\tau - \epsilon \leq t \leq \tau$  se cumple

$$|x_t^{\tau, \epsilon} - x_{\tau - \epsilon}^*| \leq \int_{\tau - \epsilon}^t |f(x_s^{\tau, \epsilon}, w)| ds \leq \epsilon C_f$$

y se tiene la misma cota para  $|x_\tau^* - x_{\tau - \epsilon}^*|$ . Por continuidad de  $D_x f$ , existe una constante  $K_w$  tal que para  $\epsilon$  suficientemente pequeño

$$|f(x_s^{\tau, \epsilon}, w) - f(x_\tau^*, w)| \leq |D_x f((1 - a)x_\tau^* + ax_s^{\tau, \epsilon})w| |x_s^{\tau, \epsilon} - x_\tau^*| \quad \text{para algún } a \in (0, 1)$$

$$\leq K |x_s^{\tau, \epsilon} - x_\tau^*|.$$

De este modo

$$\left| \frac{1}{\epsilon} \int_{\tau - \epsilon}^t f(x_s^{\tau, \epsilon}, w) - f(x_\tau^*, w) ds \right| \leq \sup_{\tau - \epsilon \leq s \leq \tau} K |x_s^{\tau, \epsilon} - x_\tau^*| \xrightarrow{\epsilon \downarrow 0} 0$$

implicando lo indicado. Defina

$$v_t := \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} (x_t^{\tau, \epsilon} - x_t^*).$$

Por el teorema 2.3.2 de [Bressan and Piccoli, 2007], para  $t \geq \tau$  el límite

$$v_t := \frac{1}{\epsilon} (x_t^{\tau, \epsilon} - x_t^*)$$

---

<sup>2</sup>Vea página 97 de [Folland, 2013] y Teorema 3.21 de dicho libro para leer la definición y enunciado de la propiedad usada en este trabajo.



existe y satisface la ecuación

$$\begin{aligned}\dot{v}_t &= D_x f(x_t^*, \theta_t^*) v_t \quad t \in [\tau, T] \\ v_\tau &= f(x_\tau^*, w) - f(x_\tau^*, \theta_\tau^*).\end{aligned}$$

La derivada en  $\tau$  es por la derecha. La minimalidad realizada por el control  $\theta^*$  implica

$$\mathbb{E}_{\mu_0} \Phi(x_T^{\tau, \epsilon}, y_0) \geq \mathbb{E}_{\mu_0} \Phi(x_T^*, y_0).$$

La suposición (A2), que  $\mu_0$  tiene soporte acotado, y que  $f$  es acotada, conllevan, por Gronwall, que existe una constante  $C$  tal que

$$|x_T^\theta| \leq C \quad \mu_0 \text{ c.s. } \forall \theta \in L^\infty([0, T], \Theta).$$

Así, junto a la continuidad de  $D_x \Phi$ , se cumple  $\frac{1}{\epsilon} [\Phi(x_T^{\tau, \epsilon}, y_0) - \Phi(x_T^*, y_0)]$  es uniformemente acotada c.s. Por convergencia dominada

$$\begin{aligned}0 &\leq \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \mathbb{E}_{\mu_0} [\Phi(x_T^{\tau, \epsilon}, y_0) - \Phi(x_T^*, y_0)] \\ &= \mathbb{E}_{\mu_0} D_x \Phi(x_T^*, y_0) v_T.\end{aligned}$$

Sea  $p^*$  la solución de la ecuación

$$\dot{p}_t^* = -(D_x f(x_s^*, \theta_s^*))^T p_s^*, \quad p_T^* = -\nabla_x \Phi(x_T^*, y_0).$$

La desigualdad anterior es

$$\mathbb{E}_{\mu_0} p_T^* v_T \leq 0.$$

En virtud de

$$\begin{aligned}\frac{d}{dt} p_t^* v_t &= \dot{p}_t^* v_t + p_t^* \dot{v}_t \quad t \in [\tau, T] \\ &= 0\end{aligned}$$

se tiene  $\mathbb{E}_{\mu_0} p_t^* v_t \leq 0, t \in [\tau, T]$ . En particular, en  $t = \tau$

$$\mathbb{E}_{\mu_0} p_\tau^* f(x_\tau^*, p_\tau^*, \theta_\tau^*) \geq \mathbb{E}_{\mu_0} p_\tau^* f(x_\tau^*, p_\tau^*, w).$$

Por la arbitrariedad de  $w$ , y  $\tau$  en el conjunto de Lebesgue de  $t \mapsto f(x_t^*, \theta_t^*)$ , que se sabe su complemento en  $[0, T]$  tiene medida 0, se tiene la conclusión del teorema.  $\square$

En la literatura de control óptimo a la ecuación de que define el proceso  $p_t$

$$\dot{p}_t^* = -\nabla_x H(x_t^*, p_t^*, \theta_t) \quad p_T^* = -\nabla_x \Phi(x_T^*, y_0)$$

se le llama ecuación adjunta. También, así como a  $x_t$  se le llama estado, a  $p_t$  se le denomina coestado.

Por otra parte, note que no hay restricciones sobre la medida de distribución  $\mu_0$  del problema teórico, en particular se puede considerar una medida soportada en una cantidad finita

de puntos, que corresponde al problema empírico. En este caso la condición de maximalidad del Principio del Máximo de Pontryagin, PMP de ahora en adelante, es

$$\frac{1}{N} \sum_{i=1}^N \nabla_{\theta} H(x_t^{i,\theta}, p^{i,\theta}, \theta_t) \geq \frac{1}{N} \sum_{i=1}^N \nabla_{\theta} H(x_t^{i,\theta}, p^{i,\theta}, w) \quad \forall w \in \Theta$$

donde  $x^{i,\theta}, p^{i,\theta}$ , son las trayectorias dadas por las respectivas ecuaciones (2.3) y (2.4) con las condiciones iniciales dadas por la muestra  $\{(x_0^i, y_0^i)\}$ .

Dado que la condición de maximalidad del PMP sólo es necesaria, se dirá que  $\theta \in L^{\infty}([0, T], \Theta)$  es una solución del PMP o que lo satisface, si los procesos  $x_t, p_t$  dados por

$$\begin{aligned} \dot{x}_t &= f(x_t, \theta_t) & x_0^* &= x_0 \\ \dot{p}_t &= -\nabla_x H(x_t, p_t, \theta_t) & p_T &= -\nabla_x \Phi(x_T, y_0) \end{aligned}$$

satisfacen

$$\mathbb{E}_{\mu_0} H(x_t, p_t, \theta_t) \geq \mathbb{E}_{\mu_0} H(x_t, p_t, w) \quad \forall w \in \Theta, \text{ c.s. } t \in [0, T].$$

Cabe mencionar que hasta cierto punto el PMP puede imaginarse como una versión infinito dimensional de multiplicadores de Lagrange. Recuerde, en multiplicadores de Lagrange se establece que si  $a^*$  es un extremo de una función  $H$  restringida a una superficie  $F(a) = 0$ , entonces cualquier dirección de perturbación de  $a$  es perpendicular al gradiente  $\nabla H$  en  $a^*$ . Por dirección de perturbación se refiere a la tangente en  $a^*$  de una curva en la superficie  $F(a) = 0$ . En la demostración del PMP se observa algo similar (tenga presente que se trabajó sobre  $L \equiv 0$ ). La restricción es en el espacio de trayectorias está dada por

$$\dot{x}_t = f(x_t, \theta_t)$$

como función del control  $\theta$ . La curva es  $\theta_t^{\tau, \epsilon}$ , y la dirección de perturbación es la derivada  $v_t$ . La ecuación

$$\frac{d}{dt} p_t^* v_t = 0 \quad E p_T v_T \leq 0$$

asemeja, en cierta manera, la condición de perpendicularidad.

Esta manera de imaginar el principio es la que se toma para idear el algoritmo numérico EMSA en el capítulo 4.

# Capítulo 3

## Relación entre controles óptimos de problemas empírico y teórico

Bajo la suposición que existen controles óptimos para los problemas empírico y teórico, interesa establecer relaciones cuantitativas entre dichos controles. Aunque no se logra dicho objetivo del todo, se demuestra el siguiente teorema.

**Teorema 3.0.1.** Bajo suposiciones adicionales (indicadas más adelante), si  $\theta^*$  satisface el PMP, entonces existen funciones  $r_1(N, s), r_2(N, s)$  y variables aleatorias  $\theta^N$ ,  $N \geq 1$ , con valores en  $L^\infty([0, T], \mathbb{R}^m)$  tales que

$$\begin{aligned}\mathbb{P}(|\theta^N - \theta^*| \geq s) &\leq r_1(N, s) \\ \mathbb{P}(F_N(\theta^N) \neq 0) &\leq r_2(N, s)\end{aligned}$$

con  $r_1(N, s), r_2(N, s) \xrightarrow{N \rightarrow \infty} 0$ , donde

$$F_N(\theta) = \frac{1}{N} \sum_{i=1}^N \nabla_{\theta} H(X_t^{i,\theta}, P_t^{i,\theta}, \theta_t)$$

y  $(X_0^i, Y_0^i)$  son iid con medida de distribución  $\mu_0$ .

Para esto se utilizará la siguiente notación. Dado  $\theta \in L^\infty([0, T], \mathbb{R}^m)$  denote por  $x^\theta, p^\theta$  los procesos dados por

$$\begin{aligned}\dot{x}_t^\theta &= f(x_t^\theta, \theta_t) & x_0^\theta &= x_0 \\ \dot{p}_t^\theta &= -\nabla_x H(x_t^\theta, p_t^\theta, \theta_t) & p_T^\theta &= -\nabla_x \Phi(x_T^\theta, y_0)\end{aligned}$$

donde la medida de distribución de  $(x_0, y_0)$  es  $\mu_0$ . Por comodidad de ahora en adelante se indicará simplemente  $(x_0, y_0) \sim \mu_0$ .

También, a lo largo de la sección se supondrá la condición

(A1') Las funciones  $f, L$  son acotadas,  $f, L, \Phi$  son dos veces continuamente diferenciables con respecto a  $x, \theta$  con derivadas acotadas y Lipschitz uniformemente.

Como se demostrará más adelante, existe una constante  $C_p$  tal que

$$|p^\theta|_{L^\infty([0,T])} \leq C_p \quad \mu_0 \text{ c.s}$$

de modo que  $\nabla_\theta H = (D_\theta f)^T p - D_\theta L$  es acotada. De esta manera, por convergencia dominada, si  $\theta^*$  es una solución del PMP, entonces para cualquier  $w \in \mathbb{R}^m$  se cumple

$$\lim_{\epsilon \rightarrow 0} \mathbb{E} \frac{H(x_t^{\theta^*}, p_t^{\theta^*}, \theta_t^* + \epsilon w) - H(x_t^{\theta^*}, p_t^{\theta^*}, \theta_t^*)}{\epsilon} = \mathbb{E} \nabla_\theta H(x_t^{\theta^*}, p_t^{\theta^*}, \theta_t^*) \cdot w.$$

La condición de maximalidad del PMP implica

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\mathbb{E} H(x_t^{\theta^*}, p_t^{\theta^*}, \theta_t^* + \epsilon w) - \mathbb{E} H(x_t^{\theta^*}, p_t^{\theta^*}, \theta_t^*)] = 0.$$

Por lo tanto

$$\mathbb{E} \nabla_\theta H(x_t^{\theta^*}, p_t^{\theta^*}, \theta_t^*) = 0.$$

Definiendo  $F : L^\infty([0, T], \mathbb{R}^m) \rightarrow L^\infty([0, T], \mathbb{R}^m)$  como

$$F(\theta) = \mathbb{E} \nabla_\theta H(x_t^\theta, p_t^\theta, \theta_t)$$

se tiene

$$F(\theta_t^*) = 0.$$

Similarmente, el control óptimo  $\theta^N$  para el problema empírico, en caso que exista, satisface

$$F_N(\theta^N) = \frac{1}{N} \sum_{i=1}^N \nabla_\theta H(x_t^{i,\theta^N}, p_t^{i,\theta^N}, \theta_t^N)$$

donde  $x_t^{i,\theta^N}, p_t^{i,\theta^N}$  son las respectivas trayectorias definidas para una muestra  $(x^i, y^i) \sim \mu_0$ .

De esta forma, una solución al problema teórico es un cero de una transformación  $F$  en un espacio de Banach, y la solución al problema empírico se piensa como una ‘aproximación aleatoria’ a dicha solución.

Antes de pasar a estudiar la relación entre los controles óptimos de los problemas empírico y teórico, se establecerán algunas propiedades respecto a los ceros de transformaciones en espacios de Banach.

### 3.1. Aproximaciones aleatorias a ceros de transformaciones de espacios de Banach.

Dado un espacio de Banach  $(U, |\cdot|)$  y un operador lineal  $T : U \rightarrow U$ , la norma del operador  $T$  se denotará  $|T|_{\text{op}}$ , es decir

$$|T|_{\text{op}} := \sup_{x \in U: |x|=1} |Tx|.$$

Además, la bola abierta centrada en un punto  $x$  de radio  $\rho$  se denotará  $B(x, \rho)$ .

**Lema 3.1.1.** (Lema de Banach) Sea  $(U, |\cdot|)$  un espacio de Banach, y  $A, B : U \rightarrow U$  operadores lineales, con  $B$  invertible, i.e.  $B^{-1}$  también es acotado, tales que

$$|A - B|_{\text{op}} \leq k \quad \text{y} \quad k|B^{-1}|_{\text{op}} < 1$$

entonces  $A$  es invertible con

$$|A^{-1}|_{\text{op}} \leq \frac{|B^{-1}|_{\text{op}}}{1 - k|B^{-1}|_{\text{op}}}$$

en particular si  $k = \frac{1}{2}|B^{-1}|_{\text{op}}^{-1}$ , entonces  $|A^{-1}|_{\text{op}} \leq 2|B^{-1}|_{\text{op}}$ .

*Demostración.* La suposición implica

$$|I - AB^{-1}|_{\text{op}} \leq k|B^{-1}|_{\text{op}} < 1$$

así  $AB^{-1}$  es invertible con inversa  $\sum_{n=0}^{\infty} (I - AB^{-1})^n$ . Como  $B$  es invertible, entonces  $A$  también lo es y

$$|A^{-1}|_{\text{op}} \leq |B^{-1}|_{\text{op}} |(AB^{-1})^{-1}| \leq \frac{|B^{-1}|_{\text{op}}}{1 - |I - AB^{-1}|} \leq \frac{|B^{-1}|_{\text{op}}}{1 - k|B^{-1}|_{\text{op}}}.$$

□

**Definición 3.1.1.** Sea  $(U, |\cdot|)$  un espacio de Banach. Una transformación  $F : X \rightarrow X$  se dirá estable en  $B(x, \rho)$  con constante  $k$  si

- $F$  es Fréchet diferenciable en  $x$ ,  $DF^{-1}(x)$  existe y es un operador lineal acotado.
- Para cualesquiera  $y, z \in B(x, \rho)$  se cumple

$$|y - z| \leq k|F(y) - F(z)|.$$

**Proposición 3.1.1.** Sea  $F : U \rightarrow U$  Fréchet diferenciable en  $x^*$ , con  $DF(x^*)$  invertible, es decir  $DF(x^*)^{-1}$  existe como operador lineal acotado. Suponga además  $DF$  es Lipschitz en  $B(x^*, \rho_0)$  con constante  $k_{DF}$ , es decir  $|DF(x) - DF(y)| \leq k_{DF}|x - y|$  para  $x, y \in B(x^*, \rho_0)$ . Entonces  $F$  es estable en  $B(x^*, \rho)$ , para  $\rho \leq \rho_0$  tal que  $\rho k_{DF}|DF(x^*)^{-1}| < \frac{1}{2}$ , y constante de estabilidad

$$k = 2|DF(x^*)^{-1}|_{\text{op}}.$$

*Demostración.* Sea  $0 < \rho \leq \rho_0$ ,  $x, y \in B(x^*, \rho)$ , entonces

$$F(x) - F(y) = R(x, y)(x - y)$$

donde

$$R(x, y) = \int_0^1 DF(sx + (1-s)y) ds.$$

Por la condición que  $DF$  es Lipschitz se cumple

$$\begin{aligned} |R(x, y) - DF(x^*)| &\leq \sup_{s \in [0,1]} |DF(sx + (1-s)y) - DF(x^*)|_{\text{op}} \\ &\leq k_{DF}. \end{aligned}$$

Si  $\rho k_{DF} |DF(x^*)^{-1}|_{\text{op}} < \frac{1}{2}$ , entonces el Lema de Banach implica  $R(x, y)$  es invertible con

$$|R(x, y)^{-1}|_{\text{op}} \leq 2|DF(x^*)^{-1}|_{\text{op}}$$

conllevando

$$|x - y| = |R(x, y)^{-1}(F(x) - F(y))| \leq 2|DF(x^*)^{-1}|_{\text{op}}|F(x) - F(y)|.$$

□

Sea  $(U, |\cdot|)$  un espacio de Banach,  $F : U \rightarrow U$  una transformación. Sea  $(\Omega, \mathcal{F}, \mathbb{P})$  un espacio de probabilidad. Una familia de transformaciones  $F_N : \Omega \times U \rightarrow U$ ,  $N \in \mathbb{N}$ , tales que  $F_N(\cdot, x)$  es medible para todo  $x$ , se dirá que son aproximaciones aleatorias de  $F$  para el cero  $x^*$  si se cumplen las siguientes condiciones

- (B1) (Estabilidad) Existe  $x^* \in U$  tal que  $F(x^*) = 0$  y  $F$  es estable en  $B(x^*, \rho)$  para algún  $\rho > 0$ .
- (B2) (Convergencia uniforme en probabilidad) Para todo  $N \geq 1$ ,  $DF(x)$  y  $DF_N(x)$  existen para todo  $x \in B(x^*, \rho)$   $\mathbb{P}$  c.s. y

$$\begin{aligned} \mathbb{P}(|F(x) - F_N(x)| \geq s) &\leq r_1(N, s) \\ \mathbb{P}(|DF(x) - DF_N(x)| \geq s) &\leq r_2(N, s) \end{aligned}$$

para algunas funciones  $r_1, r_2$  tales que  $r_1(N, s), r_2(N, s) \xrightarrow{N \rightarrow \infty} 0$  para todo  $s$ .

- (B3) (Uniformidad Lipschitz de la derivada) Existe constante  $K_{DF}$  tal que si  $x, y \in B(x^*, \rho)$  se cumple

$$|DF_N(x) - DF_N(y)|_{\text{op}} \leq K_{DF}|x - y| \quad \mathbb{P} \text{ c.s.}$$

**Teorema 3.1.1** (Aproximaciones aleatorias). Bajo las condiciones B1-B3 existen constantes  $s_0, \rho_1, C$  y variables aleatorias  $x_N \in B(x^*, \rho_1)$ , con  $\rho_1 < \rho$ , que satisfacen

$$\begin{aligned} \mathbb{P}(|x_N - x^*| \geq Cs) &\leq r_1(N, s) + r_2(N, s) \\ \mathbb{P}(F_N(x_N) \neq 0) &\leq r_1(N, s) + r_2(N, s). \end{aligned}$$

Más aún, existen conjuntos medibles  $A_N$  tales que  $\mathbb{P}(A_N) \rightarrow 1$ , y para  $\omega \in A_N$ ,  $x_N(\omega)$  es la única solución de  $F_N(x) = 0$  en  $B(x^*, \rho_1)$ . En particular  $x_N \rightarrow x^*$  y  $F_N(x_N) \rightarrow 0$  en probabilidad.

Antes de comenzar la prueba del teorema se demostrará la siguiente observación

**Observación 3.1.1.** Bajo las condiciones B1-B3, existe una constante  $s_0$  tal que para  $0 < s \leq s_0$  y  $N \geq 1$  existen conjuntos medibles  $A_N(s) \subset \Omega$  tales que

$$\begin{aligned} \mathbb{P}(A_N(s)) &\geq 1 - r_1(N, s) - r_2(N, s) \\ |F(x^*) - F_N(\omega, x^*)| &< s \quad \omega \in A_N(s). \end{aligned}$$

Además si  $0 < s_1 < s_2 \leq s_0$ , entonces  $A_N(s_1) \subset A_N(s_2)$ . Más aún

$$|DF_N(w, x^*)^{-1}|_{\text{op}} \leq 2|DF(x^*)^{-1}|$$

y  $F_N(\omega, \cdot)$  es estable en  $B(x^*, \rho_0)$  para  $\rho_0 = \min\{\rho, \frac{1}{4}(k_{DF}|DF(x^*)^{-1}|_{\text{op}})^{-1}\}$  con constante de estabilidad

$$k_{\rho_0} = 4|DF(x^*)^{-1}|.$$

*Prueba de la observación.* Defina

$$A_N(s) = \{\omega \in \Omega : |F(x^*) - F_N(\omega, x^*)| < s, |DF(x^*) - DF_N(\omega, x^*)|_{\text{op}} < s\}.$$

Para  $0 < s \leq s_0 = \frac{1}{2}|DF(x^*)^{-1}|_{\text{op}}^{-1}$ ,  $\omega \in A_N(s)$ , el lema de Banach implica

$$|DF_N(\omega, x^*)^{-1}|_{\text{op}} \leq 2|DF(x^*)^{-1}|_{\text{op}}$$

y la proposición anterior la estabilidad con la constante indicada.  $\square$

*Prueba del teorema.* Sean  $s_0, A_N(s), \rho_0$  como en la observación anterior. Defina

$$G_N(\omega, x) = x - DF_N(\omega, x^*)^{-1}F_N(\omega, x).$$

Para  $\rho_1, s$  suficientemente pequeños,  $G_N(w, \cdot)$  es una contracción en  $B(x^*, \rho_1)$  si  $w \in A_N(s)$ . A saber, para  $x, y \in B(x^*, \rho_1)$

$$\begin{aligned} G_N(\omega, x) - G_N(\omega, y) &= DF_N(\omega, x^*)^{-1} (DF_N(\omega, x^*)(x - y) - (F_N(\omega, x) - F_N(\omega, y))) \\ &= DF_N(\omega, x^*)^{-1} (DF_N(\omega, x^*)(x - y) - R_N(\omega, x, y)) (x - y) \end{aligned}$$

donde

$$R_N(\omega, x, y) = \int_0^1 DF_N(sx + (1-s)y) ds.$$

Por la propiedad B3,

$$|DF_N(\omega, x^*) - R_N(\omega, x, y)|_{\text{op}} \leq k_{DF}\rho_1, \quad \text{pues } x, y \in B(x^*, \rho_1)$$

y por la observación anterior

$$|G_N(\omega, x) - G_N(\omega, y)| \leq \alpha|x - y|, \quad \alpha = 2k_{DF}\rho_1|DF(x^*)^{-1}|_{\text{op}}.$$

Tomando  $\rho_1$  suficientemente pequeño se puede considerar  $\alpha < 1$ . Así,

$$\begin{aligned} |G_N(w, x) - x^*| &\leq |G_N(\omega, x) - G_N(\omega, x^*)| + |G_N(\omega, x^*) - x^*| \\ &\leq \alpha\rho_1 + 2|DF(x^*)^{-1}|_{\text{op}}|F_N(\omega, x) - F(x^*)| \quad \text{recuerde } F(x^*) = 0 \\ &\leq \alpha\rho_1 + 2s_0|DF(x^*)^{-1}|_{\text{op}}. \end{aligned}$$

Considerando  $s_0$  suficientemente pequeño para que  $2s_0|DF(x^*)^{-1}|_{\text{op}} < (1 - \alpha)\rho_1$  resulta que  $G_N(\omega, \cdot) : B(x^*, \rho_1) \rightarrow B(x^*, \rho_1)$  es una contracción. Por el teorema del punto fijo, para todo  $w \in A_N(s)$  existe un único  $\tilde{x}_N(\omega) \in B(x^*, \rho_1)$  tal que

$$G_N(\omega, \tilde{x}_N(\omega)) = \tilde{x}_N(\omega).$$

Además, para cualquier  $y \in B(x^*, \rho_1)$

$$\tilde{x}_N(\omega) = \lim_{k \rightarrow \infty} G_N(\omega)^{(k)}(y)$$

donde  $G_N(\omega)(y) = G_N(\omega, y)$  y el superíndice es la iteración de la composición  $k$  veces. Como  $G_N(\cdot, y)$  es medible, también los iterados  $G_N(\cdot)^{(k)}(y)$  lo son, por lo tanto

$$\omega \mapsto \tilde{x}_N(\omega) = \lim_{k \rightarrow \infty} G_N(\omega)^{(k)}(y)$$

es medible en  $A_N(s)$ . De este modo

$$x_N(\omega) = \mathbb{1}_{A_N(s)}(\omega)\tilde{x}_N(\omega) + \mathbb{1}_{A_N(s)^c}(\omega)x^*$$

es medible. Además  $A_N(s) \subset F_N(x_N) = 0$ , de modo que

$$\mathbb{P}(F_N \neq 0) \leq r_1(N, s) + r_2(N, s).$$

Por otro lado, por la estabilidad de  $F_N$  se tiene, recordando que para  $\omega \in A_N(s)$  se cumple  $F_N(\omega, x_N(s)) = F(x^*) = 0$ ,

$$\begin{aligned} |x_N(\omega) - x^*| &\leq k_{\rho_1} |F_N(\omega, x_N(\omega)) - F_N(\omega, x^*)| \\ &\leq 4|DF(x^*)^{-1}|_{\text{op}} |F(x^*) - F_N(\omega, x^*)| \\ &\leq 4s|DF(x^*)^{-1}|_{\text{op}}. \end{aligned}$$

Escogiendo  $C = 4|DF(x^*)^{-1}|_{\text{op}}$  se satisface

$$\mathbb{P}(|x_N(\omega) - x^*| \geq Cs) \leq r_1(N, s) + r_2(N, s)$$

Dada la contención  $A_N(s_1) \subset A_N(s_2)$  para  $s_1 \leq s_2$ , y la unicidad de  $\tilde{x}_N(\omega)$ , los conjuntos  $A_N$  que se mencionan en el enunciado pueden tomarse como  $A_N(s_0)$ .  $\square$

## 3.2. Verificación de condiciones

Tomando en cuenta el teorema anterior, se intentará verificar las condiciones para probar el teorema enunciado al principio de la sección.

**Propiedad 3.2.1.** Existen constantes  $C, k$  tales que para todo  $\theta, \phi \in L^\infty([0, T], \mathbb{R}^m)$

$$\begin{aligned} |x^\theta|_{L^\infty} + |p^\theta|_{L^\infty} &\leq C \\ |x^\theta - x^\phi|_{L^\infty} + |p^\theta - p^\phi|_{L^\infty} &\leq k|\theta - \phi|_{L^\infty}. \end{aligned}$$

*Demostración.* Como  $f$  es acotada, entonces

$$|x_t^\theta| \leq |x_0| + \int_0^t |f(x_s^\theta, \theta_s)| ds$$



es acotada independientemente de  $\theta$  c.s., pues  $\mu_0$  tiene soporte acotado. En virtud que las parciales de  $f, L, \Phi$  son acotadas, entonces de

$$|p_t^\theta| \leq |\nabla_x \Phi(x_T^\theta, y_0)| + \int_t^T |D_x f(x_s^\theta, \theta_s)^T| |p_t^\theta| + |\nabla_x L(x_s^\theta, \theta_s)| ds$$

se sigue por la desigualdad de Gronwall

$$|p_t^\theta| \leq [C_{\nabla_x \Phi} + C_{\nabla_x L}] \exp\{C_{D_x f} T\}.$$

Lo que implica la primera desigualdad.

Para la segunda, por Gronwall

$$\begin{aligned} |x_t^\theta - x_t^\phi| &= \left| \int_0^t f(x_s^\theta, \theta_s) - f(x_s^\phi, \phi_s) ds \right| \\ &= \left| \int_0^t \int_0^1 Df(rx_s^\phi + (1-r)x_s^\theta, r\phi_s + (1-r)\theta_s)(x_s^\theta - x_s^\phi, \theta_s - \phi_s) \right| dr ds \\ &\leq C \int_0^t |x_s^\theta - x_s^\phi| ds + CT|\theta - \phi|_{L^\infty} \\ &\leq CT \exp\{CT\} |\theta - \phi|_{L^\infty}. \end{aligned}$$

Similarmente

$$|p_t^\theta - p_t^\phi| \leq |\nabla_x \Phi(x_T^\theta, y_0) - \nabla_x \Phi(x_T^\phi, y_0)| + \left| \int_t^T \nabla_x H(x_s^\theta, p_s^\theta, \theta_s) - \nabla_x H(x_s^\phi, p_s^\phi, \phi_s) ds \right|.$$

Como  $\nabla_x H(x, p, w) = D_x f(x, w)^T p - \nabla_x L(x, w)$ , y  $p$  es acotada c.s., entonces  $D\nabla_x H$  es acotada (c.s.). Luego, por Gronwall

$$\begin{aligned} |p_t^\theta - p_t^\phi| &\leq C_{D\nabla_x \Phi} |x_T^\theta - x_T^\phi| + \int_t^T C(|x_s^\theta - x_s^\phi| + |p_s^\theta - p_s^\phi| + |\theta_s - \phi_s|) ds \\ &\leq (k_1 + TC + k_2) |\theta - \phi|_{L^\infty} \exp\{CT\}. \end{aligned}$$

Que implica la segunda desigualdad. □

Ahora considere  $x^\theta$  como una transformación  $\theta \mapsto x^\theta$  de  $L^\infty([0, T], \mathbb{R}^m)$  en si mismo, similarmente  $p^\theta$ .

**Propiedad 3.2.2.** Existen constantes  $C, k$  tales que para todo  $\theta, \phi \in L^\infty([0, T], \mathbb{R}^m)$ .

$$\begin{aligned} |Dx^\theta|_{\text{op}} + |Dp^\theta|_{\text{op}} &\leq C \\ |Dx^\theta - Dx^\phi|_{\text{op}} + |Dp^\theta - Dp^\phi|_{\text{op}} &\leq k|\theta - \phi|_{L^\infty}. \end{aligned}$$

*Demostración.* Se demostrará primero que para todo  $\eta \in L^\infty([0, T], \mathbb{R}^m)$  con  $|\eta|_{L^\infty} = 1$  se tiene que

$$[(Dx^\theta)\eta]_t := \lim_{\epsilon \rightarrow 0} \frac{x_t^{\theta+\epsilon\eta} - x_t^\theta}{\epsilon}$$

existe y satisface la ecuación diferencial

$$\frac{d}{dt}[(Dx^\theta)\eta]_t = (D_x f_t^\theta)[(Dx^\theta)\eta]_t + (D_\theta f_t^\theta) [(Dx^\theta)\eta]_0 = 0$$

donde  $f_t^\theta = f(x_t^\theta, \theta_t)$ .

Sea  $\eta$  como se indicó y  $\delta_\eta x_t^\theta$  la solución en  $[0, T]$  de

$$\frac{d}{dt}\delta_\eta x_t^\theta = D_x f_t^\theta(\delta_\eta x_t^\theta) + D_\theta f_t^\theta \eta_t \quad \delta_\eta x_0^\theta = 0.$$

Defina  $y_t^\epsilon = x_t^\theta + \epsilon \delta x_t^\theta$ . Note que se desea demostrar

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (x_t^{\theta+\epsilon\eta} - y_t^\epsilon) = 0 \quad \text{en } L^\infty([0, T], \mathbb{R}^m).$$

Recuerde que  $x_t^{\theta+\epsilon\eta}$  es el único punto fijo de

$$q_t \mapsto x_0 + \int_0^t f(q_s, (\theta + \epsilon\eta)_s) ds =: \varphi(q, (\theta + \epsilon\eta))$$

con  $\varphi : C([0, T], \mathbb{R}^m) \rightarrow C([0, T], \mathbb{R}^m)$ , entonces por 1.2.2

$$\left| \frac{1}{\epsilon} (x_t^{\theta+\epsilon\eta} - y_t^\epsilon) \right|_{\dagger} \leq \frac{2}{|\epsilon|} |\varphi(y_t^\epsilon, (\theta + \epsilon\eta)) - y_t^\epsilon|_{\dagger}.$$

Como  $|\cdot|_{\dagger}$  y  $|\cdot|_{L^\infty}$  son normas equivalentes en  $C([0, T], \mathbb{R}^m)$ , será suficiente demostrar

$$\limsup_{\epsilon \rightarrow 0} \left| \frac{1}{\epsilon} (\varphi(y_t^\epsilon, (\theta + \epsilon\eta)) - y_t^\epsilon) \right| = 0.$$

Recordando que por notación  $y_s^\epsilon - y_s^0 = \epsilon \delta_\eta x_s^\theta$ ,  $x_t^\theta = x_0 + \int_0^t f(x_s^\theta, \theta_s) ds$ ,  $\delta_\eta x_t^\theta = \int_0^t (D_x f_s^\theta)[\delta_\eta x_s^\theta] + (D_\theta f_s^\theta)\eta_s ds$  se tiene

$$\begin{aligned} \left| \frac{1}{\epsilon} [\varphi(y_t^\epsilon, (\theta + \epsilon\eta)_t) - y_t^\epsilon] \right| &= \left| \frac{1}{\epsilon} [x_0 + \int_0^t f(y_s^\epsilon, (\theta + \epsilon\eta)_s) ds - y_t^\epsilon] \right| \\ &= \left| [x_0 + \int_0^t f(y_s^0, \theta_s) ds \right. \\ &\quad + \int_0^t D_x f_s^\theta (y_s^\epsilon - y_s^0) ds + \int_0^t D_\theta f_s^\theta (\epsilon\eta)_s ds \\ &\quad + \int_0^t \int_0^1 [D_x f(a, \theta, \eta, \epsilon, s) - D_x f_s^\theta](y_s^\epsilon - y_s^0) dad s \\ &\quad + \int_0^t \int_0^1 [D_\theta f(a, \theta, \eta, \epsilon, s) - D_\theta f_s^\theta](\epsilon\eta)_s dad s \\ &\quad \left. - x_s^\theta - \epsilon \delta_\eta x_s^\theta] \right| \\ &\leq \int_0^t \int_0^1 |(D_x f(a, \theta, \eta, \epsilon, s) - D_x f_s^\theta)(\delta_\eta x_s^\theta)| dad s \\ &\quad + \int_0^t \int_0^1 |(D_\theta f(a, \theta, \eta, \epsilon, s) - D_\theta f_s^\theta)\eta_s| dad s. \end{aligned}$$

Por la ecuación diferencial que define a  $\delta_\eta x_t^\theta$ ,  $\delta_\eta x_t^\theta$  es uniformemente acotada en  $[0, T]$  ( $\mu_0$  c.s.), por ejemplo use Gronwall. De ahí, usando que  $D_x f$  es Lipschitz y en particular uniformemente continua, se cumple

$$D_x f(a, \theta, \eta, \epsilon, s) := D_x f((1-a)x_s^\theta + a(x_s^\theta + \epsilon \delta_\eta x_s^\theta), (1-a)\theta_s + a(\theta + \epsilon \eta)_s) \\ \xrightarrow{\epsilon \rightarrow 0} D_x f_s^\theta \quad \text{uniformemente.}$$

Por lo tanto

$$\left| \frac{1}{\epsilon} (\varphi(y_t^\epsilon, (\theta + \epsilon \eta)_t) - y_t^\epsilon) \right|_{L^\infty} \rightarrow 0$$

que es lo que se quería demostrar.

Por otra parte

$$\eta \mapsto (Dx^\theta)\eta$$

es lineal, la aditividad es por la unicidad de la solución de ecuaciones diferenciales. Finalmente, como  $D_x f$  es acotada, por Gronwall existe una constante  $C$ , que depende de  $T$ , tal que

$$|(Dx^\theta)\eta|_{L^\infty} \leq C$$

por lo que  $\theta \mapsto x^\theta$  es derivable.

Ahora bien, para  $\eta$  como antes,

$$\begin{aligned} |[(Dx^\theta)\eta]_t - [(Dx^\phi)\eta]_t| &\leq \int_0^t |D_x f_s^\theta| |(Dx^\theta)\eta_s - (Dx^\phi)\eta_s| ds \\ &+ \int_0^t |D_x f_s^\theta - D_x f_s^\phi| |(Dx^\phi)\eta_s| ds \\ &+ \int_0^t |D_x f_s^\theta - D_\theta f_s^\phi| |\eta_s| ds. \end{aligned}$$

Por la propiedad anterior 3.2.1

$$\begin{aligned} |D_x f_s^\theta - D_x f_s^\phi| &\leq k|\theta - \phi|_{L^\infty} \\ |D_\theta f_s^\theta - D_\theta f_s^\phi| &\leq k|\theta - \phi|_{L^\infty} \end{aligned}$$

y por Gronwall

$$|(Dx^\theta)\eta - (Dx^\phi)\eta| \leq k|\theta - \phi|_{L^\infty}.$$

Similarmente se demuestra para  $p^\theta$ , primero probando que para  $\eta \in L^\infty([0, T], \mathbb{R}^n)$ ,  $|\eta|_{L^\infty} = 1$ , se cumple

$$\begin{aligned} [(Dp^\theta)\eta]_T &= -D_{x,x}^2 \Phi(x_T^\theta, y_0) [(Dx^\theta)\eta]_T \\ \frac{d}{dt} [(Dp^\theta)\eta]_t &= -D_{x,x}^2 H_t^\theta [(Dx^\theta)\eta]_t - D_{p,x}^2 [(Dp^\theta)\eta]_t - D_{\theta,x}^2 H_t^\theta \eta_t \end{aligned}$$

donde

$$(Dp^\theta)\eta := \lim_{\epsilon \rightarrow 0} \frac{p^{\theta+\epsilon\eta} - p^\theta}{\epsilon}.$$

Note que la condición en  $T$  es sólo regla de la cadena y es la primera que se obtiene.  $\square$

**Propiedad 3.2.3.** Sea  $g : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  tal que para cada compacto  $R \subset \mathbb{R}^d \times \mathbb{R}^d$ , existen constantes  $C_R, k_R$  tales que

$$\begin{aligned} |Dg(x, p, w)| &\leq C_R \quad \forall (x, p, w) \in R \times \mathbb{R}^m \\ |Dg(x_1, p_1, w_1) - Dg(x_2, p_2, w_2)| &\leq k[|x_1 - x_2| + |p_1 - p_2| + |w_1 - w_2|] \end{aligned}$$

con la segunda desigualdad válida para todo  $(x_1, p_1, w_1), (x_2, p_2, w_2) \in R \times \mathbb{R}^m$ .

Defina  $G : L^\infty([0, T], \mathbb{R}^m) \rightarrow L^\infty([0, T], \mathbb{R}^m)$  mediante

$$G(\theta)_t = g(x^\theta, p_t^\theta, \theta_t).$$

Entonces  $G$  es diferenciable,  $DG$  es acotada, y Lipschitz  $\mu_0$  c.s. Esto es, existen constantes  $C, k$  que cumplen

$$\begin{aligned} |DG(\theta)| &\leq C \quad \mu_0 \quad \text{c.s.} \\ |DG(\theta) - DG(\phi)| &\leq k|\theta - \phi|_{L^\infty} \quad \mu_0 \quad \text{c.s.} \end{aligned}$$

*Demostración.* Sea  $\eta \in L^\infty([0, T], \mathbb{R}^m)$ ,  $|\eta| \leq 1$ . Entonces

$$\begin{aligned} \left[ \lim_{\epsilon \rightarrow 0} \frac{G(\theta + \epsilon\eta) - G(\theta)}{\epsilon} \right]_t &= Dg_t^\theta [D_\theta(x^\theta, p^\theta, \theta)\eta]_t \quad \text{en } L^\infty \\ &= (D_x g_t^\theta)[(Dx^\theta)\eta]_t + (D_p g_t^\theta)[(Dp^\theta)\eta]_t + (D_\theta g_t^\theta)\eta_t. \end{aligned}$$

A saber, denotando  $(x_t^{\theta+\epsilon\eta} - x_t^\theta, p_t^{\theta+\epsilon\eta} - p_t^\theta, \epsilon\eta_t)$  por  $(\Delta(x, p, \theta))_t \Big|_\theta^{\theta+\epsilon\eta}$  se cumple

$$\begin{aligned} &\left| \frac{1}{\epsilon} \left[ g(x_t^{\theta+\epsilon\eta}, p_t^{\theta+\epsilon\eta}, (\theta + \epsilon\eta)_t) - g(x_t^\theta, p_t^\theta, \theta_t) - \epsilon Dg_t^\theta [D_\theta(x^\theta, p^\theta, \theta)\eta]_t \right] \right| \\ &\leq \left| \frac{1}{\epsilon} \left[ \int_0^1 Dg(a, \theta, \eta, \epsilon, t) (\Delta(x, p, \theta))_t \Big|_\theta^{\theta+\epsilon\eta} da - \epsilon Dg_t^\theta [D_\theta(x^\theta, p^\theta, \theta)\eta]_t \right] \right| \\ &\leq \left| \int_0^1 [Dg(a, \theta, \eta, \epsilon, t) - Dg_t^\theta] \frac{\Delta(x, p, \theta)_t \Big|_\theta^{\theta+\epsilon\eta}}{\epsilon} da \right| \\ &\quad + \left| Dg_t^\theta \left( \frac{\Delta(x, p, \theta)_t \Big|_\theta^{\theta+\epsilon\eta}}{\epsilon} - \epsilon [D_\theta(x^\theta, p^\theta, \theta)\eta]_t \right) \right|. \end{aligned}$$

Por la propiedad anterior 3.2.1, existe  $C$  tal que

$$|(x^\phi, p^\phi)| \leq C \quad \forall \phi \in L^\infty([0, T], \mathbb{R}^m) \quad \mu_0 \text{ c.s.}$$

Por las condiciones en  $g$  existe  $k$  independiente de  $\theta$  que satisface

$$[Dg(a, \theta, \eta, \epsilon, t) - Dg_t^\theta]_{L^\infty} \leq k \left| \Delta(x, p, \theta) \Big|_\theta^{\theta+\epsilon\eta} \right|_{L^\infty} \xrightarrow{\epsilon \rightarrow 0} 0.$$

En virtud de

$$\frac{\Delta(x, p, \theta) \Big|_\theta^{\theta+\epsilon\eta}}{\epsilon} \rightarrow D_\theta(x^\theta, p^\theta, \theta)\eta \quad \text{en } L^\infty$$

pues  $x^\theta, p^\theta$  son Fréchet diferenciables, para  $\epsilon$  suficientemente pequeño,  $\frac{\Delta(x,p,\theta)}{\epsilon} \Big|_\theta^{\theta+\epsilon\eta}$  es acotado. Así,

$$\left| \int_0^1 [Dg(a, \theta, \eta, \epsilon, t) - Dg_t^\theta] \frac{\Delta(x, p, \theta)_t}{\epsilon} \Big|_\theta^{\theta+\epsilon\eta} da \right|_{L^\infty} \xrightarrow{\epsilon \rightarrow 0} 0.$$

Al ser  $Dh_t^\theta$  acotado, se cumple

$$\left| Dg_t^\theta \left( \frac{\Delta(x, p, \theta)_t}{\epsilon} \Big|_\theta^{\theta+\epsilon\eta} - \epsilon [D_\theta(x^\theta, p^\theta, \theta)\eta]_t \right) \right|_{L^\infty} \xrightarrow{\epsilon \rightarrow 0} 0.$$

Que implica  $G$  es (Fréchet) diferenciable. Debido a que  $Dg_t^\theta$  es acotada uniformemente, es decir existe  $C$  tal que  $|Dg_t^\theta| \leq C$  para todo  $\theta \in L^\infty([0, T], \mathbb{R}^m)$ , y que  $Dx^\theta, Dp^\theta$  también lo son, existe una constante  $C$  la cual satisface

$$DG(\theta) \leq C.$$

Finalmente, para  $\theta, \phi \in L^\infty([0, T], \mathbb{R}^m)$ ,  $\eta$  como antes, se cumple

$$\begin{aligned} |DG(\theta)\eta_t - DG(\phi)\eta_t| &\leq |D_x g_t^\theta [(Dx^\theta)\eta]_t - D_x g_t^\phi [(Dx^\phi)\eta]_t| \\ &\quad + |D_p g_t^\theta [(Dp^\theta)\eta]_t - D_p g_t^\phi [(Dp^\phi)\eta]_t| \\ &\quad + |D_\theta g_t^\theta \eta_t - D_\theta g_t^\phi \eta_t|. \end{aligned}$$

Note además

$$\begin{aligned} |D_x g_t^\theta [(Dx^\theta)\eta]_t - D_x g_t^\phi [(Dx^\phi)\eta]_t| &\leq |(D_x g_t^\theta - D_x g_t^\phi) [(Dx^\theta)\eta]_t| \\ &\quad + |D_x g_t^\phi ([(Dx^\theta)\eta]_t - [(Dx^\phi)\eta]_t)| \\ &\leq |(D_x g_t^\theta - D_x g_t^\phi) C_{Dx} + C_{D_x g} k |\theta - \phi|_{L^\infty}| \\ &\leq C (k [|x^\theta - x^\phi]_{L^\infty} + |p^\theta - p^\phi|_{L^\infty} + |\theta - \phi|_{L^\infty}) \\ &\quad + |\theta - \phi|_{L^\infty} \\ &\leq k |\theta - \phi|_{L^\infty}. \end{aligned}$$

Similarmente para los otros términos, el de  $p^\theta$  y el de  $\theta$ , implicando que  $DG$  es Lipschitz.  $\square$

Por otra parte, observe que si  $g : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$ ,  $(x, p, w) \mapsto h(x, p, w)$  es Lipschitz y acotada, entonces para  $|p| \leq C$ ,  $hp$  es Lipschitz en  $\mathbb{R}^d \times B(0, C) \times \mathbb{R}^m$ . A saber, sean  $(x_i, p_i, w_i) \in \mathbb{R}^d \times B(0, C) \times \mathbb{R}^m$ ,  $i = 1, 2$ , denotando por  $h_i = h(x_i, p_i, w_i)$ , se tiene

$$\begin{aligned} |h_1 p_1 - h_2 p_2| &\leq |h_1| |p_1 - p_2| + |h_1 - h_2| |p_2| \\ &\leq C |p_1 - p_2| + k [|x_1 - x_2| + |p_1 - p_2| + |w_1 - w_2|] \\ &\leq k [|x_1 - x_2| + |p_1 - p_2| + |w_1 - w_2|]. \end{aligned}$$

De esta manera  $\nabla_\theta H = (D_\theta f)^T p - \nabla_\theta L$  satisface las condiciones de la propiedad anterior, ya que

$$\begin{aligned} D_{x,\theta} H &= D_x(D_\theta f)^T p - D_{x,\theta} L \\ D_{p,\theta} H &= (D_\theta f)^T \\ D_\theta^2 H &= D_\theta(D_\theta f)^T p - D_\theta^2 L. \end{aligned}$$

Así, por las últimas tres proposiciones, las funciones

$$F_N(\theta) = \frac{1}{N} \sum_{i=1}^N \nabla_\theta H(X^{i,\theta}, P^{i,\theta}, \theta_t)$$

satisfacen la condición B3 para ser aproximaciones aleatorias, es decir, uniformemente Lipschitz.

Para ver que  $F_N(\cdot, x) : \Omega \rightarrow L^\infty([0, T], \mathbb{R}^m)$  es medible, será suficiente mostrar  $\omega \mapsto x^\theta(\omega)$  y  $\omega \mapsto p^\theta(\omega)$  son medibles, pues  $(x, p, \theta) \mapsto \nabla_\theta H(x, p, \theta)$  definida en  $L^\infty \times L^\infty \times L^\infty \rightarrow L^\infty$  es continua. De hecho, por el comentario descrito después de la prueba de la última proposición, se sabe que es localmente Lipschitz.

Para ver que  $\omega \mapsto x^\theta(\omega)$  es medible, se usará  $x^\theta(\omega) \in C([0, T], \mathbb{R}^m)$ , pues es una solución clásica de un ecuación diferencial ordinaria, para todo  $\omega \in \Omega$ . Así, es suficiente mostrar que para todo  $h \in C([0, T], \mathbb{R}^m)$ ,  $\rho_0 > 0$ , el conjunto

$$\{\omega : x^\theta(\omega) \in B(h, \rho_0)\}$$

es medible.

En virtud de que  $f$  es acotada y Lipschitz, existe flujo  $\varphi$  para la ecuación diferencial

$$\dot{x}_t = f(x_t, \theta_t)$$

de modo que  $X_t^\theta = \varphi(X_0, t)$  es medible para todo  $t$ . Por otra parte,  $x^\theta(\omega) \in B(h, \rho_0)$  implica que existe  $\rho < \rho_0$  tal que  $|x^\theta(\omega) - h|_{L^\infty} = \rho$ , conllevando

$$\omega \in \bigcap_{t \in \mathbb{Q} \cap [0, T]} \{|X_t^\theta - h_t| \leq \rho\} \in \mathcal{F}.$$

Recuerde  $\mathcal{F}$  es la  $\sigma$ -álgebra del espacio de probabilidad que se fijó. Viceversa, por continuidad, dado  $0 < \rho < \rho_0$ ,

$$\omega \in \bigcap_{t \in \mathbb{Q} \cap [0, T]} \{|X_t^\theta - h_t| \leq \rho\}$$

implica  $x^\theta(\omega) \in B(h, \rho)$ . Esto es,

$$\{\omega : x^\theta(\omega) \in B(h, \rho_0)\} = \bigcup_{\rho \in \mathbb{Q} \cap (0, \rho_0)} \bigcap_{t \in \mathbb{Q} \cap [0, T]} |X_t^\theta - h_t| \leq \rho$$

que es medible. Similarmente para  $p^\theta$ , con la diferencia que  $\dot{p}^\theta = -(D_x f)^T p + \nabla_x L$ , al ser una ecuación lineal no homogénea con  $D_x f, D_x L$  acotadas, el flujo para  $p$  existe.

Para verificar la condición B2 sobre convergencia uniforme en probabilidad se usará la siguiente versión de la desigualdad de Hoeffding, encontrada como corolario 2 en [Pinelis and Sakhnenko, 1986]

**Teorema 3.2.1** (Desigualdad de Hoeffding en espacios de Banach). Sean  $X_1, X_2, \dots$  variables aleatorias, independientes, con valores en un espacio de Banach separable  $(\mathfrak{X}, |\cdot|)$ , y  $S_n = X_1 + \dots + X_n$ . Suponga que existen  $x_i \in \mathfrak{X}$ ,  $y \in \mathbb{R}$  tales que  $\mathbb{P}(|X_i - x_i| \leq y) = 1$ , para todo  $i$ , entonces

$$\mathbb{P}(|S_n| - \mathbb{E}|S_n| \geq x) \leq \exp \left\{ \frac{-x^2}{2B_n^2 + \frac{2}{3}xy} \right\}$$

donde  $B_n^2 = \sum_1^n \mathbb{E}|X_i - x_i|^2$ .

**Observación 3.2.1.** Las funciones

$$\begin{aligned} F(\theta) &= \mathbb{E}_{\mu_0} \nabla_{\theta} H(x^{\theta}, p^{\theta}, \theta) \\ F_n(\theta) &= \frac{1}{N} \sum_{i=1}^N \nabla_{\theta} H(x^{i,\theta}, p^{i,\theta}, \theta) \end{aligned}$$

donde  $(x_0^i, y_0^i) \sim \mu_0$  son iid, satisfacen la condición B2 de convergencia uniforme en probabilidad.

*Demostración.* La diferenciabilidad de  $\theta \mapsto \nabla_{\theta} H(X^{\theta}, P^{\theta}, \theta)$  es por la propiedad 3.2.3, además la derivada es acotada. Así, por convergencia dominada,  $F$  es derivable y

$$DF(\theta) = \mathbb{E}_{\mu_0} D_{\theta} \nabla_{\theta} H(X^{\theta}, P^{\theta}, \theta).$$

Aquí, el gradiente  $\nabla_{\theta}$  es respecto a la tercer variable de  $H$ , la del valor del control. La derivada  $D_{\theta}$ , es en  $L^{\infty}([0, T], \mathbb{R}^m)$ , con  $\theta$  como control.

Por otra parte,

$$F_N - F = \sum_{i=1}^N \frac{1}{N} [\nabla_{\theta} H(X^{i,\theta}, P^{i,\theta}, \theta) - \mathbb{E} \nabla_{\theta} H(X^{\theta}, P^{\theta}, \theta)].$$

Así, por la desigualdad de Hoeffding, existen constantes  $C_1, C_2$  tales que

$$\mathbb{P}(|F_N - F| - \mathbb{E}|F_N - F| \geq x) \leq \exp \left\{ \frac{-Nx^2}{2C_1 + \frac{2}{3}C_2x} \right\}.$$

Esto resulta de tomar  $x_i = 0$ ,  $y$  una cota de  $\frac{1}{N} |\nabla_{\theta} H(X^{i,\theta}, P^{i,\theta}, \theta) - \mathbb{E} \nabla_{\theta} H(X^{\theta}, P^{\theta}, \theta)|$ , la cual se puede tomar independiente de  $\theta$  pues  $\nabla_{\theta} H(x^{\theta}, P^{\theta}, \theta)$  es acotada uniformemente en  $\theta$   $\mu_0$  c.s., y notar que

$$\begin{aligned} B_N^2 &= \sum_{i=1}^N \mathbb{E} \left| \frac{1}{N} [\nabla_{\theta} H(X^{i,\theta}, P^{i,\theta}, \theta) - \mathbb{E} \nabla_{\theta} H(X^{\theta}, P^{\theta}, \theta)] \right|^2 \\ &= \frac{1}{N} \mathbb{E} |[\nabla_{\theta} H(X^{i,\theta}, P^{i,\theta}, \theta) - \mathbb{E} \nabla_{\theta} H(X^{\theta}, P^{\theta}, \theta)]|^2 \quad \text{pues } (X_0^i, Y_0^i) \text{ son iid} \\ &= \frac{C_1}{N}. \end{aligned}$$

Por el corolario 7.10 de [Talagrand and Ledoux, 1991],

$$F_N(\theta) \rightarrow F(\theta) \quad \text{c.s.}$$

En virtud de que  $\nabla_{\theta}H(X^{\theta}, P^{\theta}, \theta)$  es acotada, se tiene  $\mathbb{E}|F_N - F| \xrightarrow[N \rightarrow 0]{} 0$ . Así, dado  $s > 0$ , defina

$$N(s) = \inf\{N : \mathbb{E}|F_N - F| \leq \frac{s}{2}\}$$

y

$$r_1(N, s) = \begin{cases} 1 & \text{si } N < N(s) \\ \exp\left\{\frac{-N\left(\frac{s}{2}\right)^2}{2C_1 + \frac{1}{3}C_2x}\right\} & \text{para } N \geq N(s). \end{cases}$$

Entonces por construcción

$$\mathbb{P}(|F_N - F| \geq s) \leq r_1(N, s)$$

y  $r_1(N, s) \rightarrow 0$  cuando  $N \rightarrow \infty$ . Similarmente para  $DF_N, DF$  existe  $r_2(N, s)$  que satisface

$$\mathbb{P}(|DF_N - DF| \geq s) \leq r_2(N, s)$$

y  $r_2(N, s) \rightarrow 0$  si  $N \rightarrow \infty$ . □

De esta manera, por el teorema 3.1.1 se tiene el teorema enunciado al principio de la sección.

**Teorema 3.2.2** (Consistencia de la solución al problema empírico). Sea  $\theta^*$  una solución estable de

$$\mathbb{E}_{\mu_0} \nabla_{\theta}H(X^{\theta}, P^{\theta}, \theta) = 0$$

en  $B(\theta^*, \rho_0)$ , entonces existen constantes positivas  $\rho < \rho_0, C, s_0$ , funciones  $r_1(N, s), r_2(N, s)$ , y variables aleatorias  $\theta^N \in B(\theta^*, \rho_1) \subset L^{\infty}([0, T], \mathbb{R}^m)$  tales que

$$\begin{aligned} \mathbb{P}(|\theta^* - \theta^N| \geq Cs) &\leq r_1(N, s) \\ \mathbb{P}(|F_N(\theta^N)| \neq 0) &\leq r_2(N, s) \end{aligned}$$

con  $r_1(N, s), r_2(N, s) \rightarrow 0$ , cuando  $N \rightarrow \infty$ . Más aún existen conjuntos medibles  $A_N$  tales que  $\mathbb{P}(A_N) \rightarrow 1$ , y para  $\omega \in A_N$ ,  $\theta^N(\omega)$  es la única solución en  $B(\theta^*, \rho)$  de  $F_N(\theta) = 0$ .

El teorema anterior tiene varias consecuencias de carácter teórico y práctico. Las consecuencias teóricas inmediatas son la consistencia y robustez de la solución del problema empírico. Desde la perspectiva práctica se observa que si el problema empírico admite un control óptimo, entonces, por la unicidad de la solución  $F_N(\theta) = 0$  en los conjuntos  $A_N$  del teorema, con alta probabilidad  $\theta^N$  será igual al control óptimo del problema empírico. Así, independientemente de la manera de hallar el control óptimo, con alta probabilidad, éste será cercano al control ideal, el del problema teórico. En particular, aún cuando cambie la muestra de entrenamiento, el control hallado por entrenamiento no cambiará mucho. Por otra parte, aún cuando no exista control óptimo para el problema empírico, se puede usar como control la solución de  $F_N(\theta) = 0$ . Por lo establecido en el teorema, con alta probabilidad, éste tendrá propiedades similares al que tendría uno óptimo, y además se puede garantizar que, con alta



---

probabilidad, no sólo aproximará bien en la muestra de entrenamiento, sino también en una de prueba.

Finalmente, el teorema establece la garantía que cuando la muestra es grande, se puede confiar en obtener pesos tan eficientes como los más eficientes posibles, i.e. los del problema teórico.



# Capítulo 4

## Algoritmo EMSA

Como se mencionó en la introducción, inspirados en el Principio del Máximo de Pontryagin, los investigadores Li, E, Chen y Tai propusieron en [Li et al., 2017] un nuevo algoritmo, nombrado EMSA, para encontrar los pesos de una red residual que aproxima una función desconocida de interés. Dicho método busca el control óptimo del problema empírico asociado, ecuación (2.1), y usa que para encontrarlo numéricamente hay que discretizar, lo cual automáticamente da los pesos para la red residual asociada. Una de las propiedades más importante del algoritmo EMSA es que bajo condiciones generales siempre converge. La demostración de este hecho se basa en el lema 2 del artículo citado. Aunque todas las demostraciones en [Li et al., 2017] son completas, es decir los hechos matemáticos son ciertos, la prueba del lema 2 podría beneficiarse de una explicación más detallada. Así, parte del propósito de este capítulo es presentar el algoritmo EMSA, junto a sus propiedades más importantes, y detallando únicamente la demostración del lema 2 del artículo.

A lo largo del capítulo se seguirá trabajando bajo la suposición A1' del capítulo 3, es decir

**(A1')** Las funciones  $f, L$  son acotadas, además  $f, L, \Phi$  son dos veces continuamente diferenciables con derivadas acotadas y Lipschitz.

Por otra parte, para evitar la aglomeración de símbolos, el producto interno en  $\mathbb{R}^n$  se denota con punto  $\cdot$  cuando se quiera enfatizar, o simplemente escribiendo juntos los objetos.

Para redes neuronales lo usual es que el término de regularización,  $L$ , sólo sea una función del peso  $\theta \in \mathbb{R}^m$ , es decir de la forma  $L(\theta)$ . De este modo, se supondrá que  $L$  no depende del estado  $x$ . Por otra parte, para el problema de entrenamiento la muestra  $\{(x^i, y^i)\}_{i=1}^N$  es fija. Por consiguiente, bajo la notación

$$\begin{aligned} X_0 &= (x_0^1, x_0^2, \dots, x_0^N) \\ X_t^\theta &= (x_t^1, x_t^2, \dots, x_t^N) \\ f(X_t^\theta, \theta_t) &= (f(x_t^{1,\theta}, \theta_t), f(x_t^{2,\theta}, \theta_t), \dots, f(x_t^{N,\theta}, \theta_t)) \\ \Phi(X_T^\theta) &= \sum_{n=1}^N \frac{1}{N} \Phi(x_T^{n,\theta}, y^n) \end{aligned}$$

es posible reescribir el problema empírico en la forma

$$\begin{aligned} \inf_{\theta \in L^\infty([0,T],\Theta)} J(\theta) &= \Phi(X_T^\theta) + \int_0^T L(\theta_s) ds \\ \text{sujeto a } \dot{X}_t^\theta &= f(X_t^\theta, \theta_t) \quad X_0^\theta = X_0. \end{aligned}$$

El Hamiltoniano se expresa como

$$H(x, p, \theta) = p \cdot f(x, \theta) - L(\theta)$$

y el Principio del Máximo de Pontryagin, para el problema empírico, toma la forma: *Si  $\theta$  es control óptimo para el problema empírico, entonces los procesos  $X_t^\theta, P_t^\theta$  dados por*

$$\begin{aligned} \dot{X}_t^\theta &= f(X_t^\theta, \theta_t) \quad X_0^\theta = X_0 \\ \dot{P}_t^\theta &= -\nabla_x H(X_t^\theta, P_t^\theta, \theta_t) \quad P_T^\theta = -\nabla \Phi(X_T^\theta) \end{aligned}$$

satisfacen

$$H(X_t^\theta, P_t^\theta, \theta_t) \geq H(X_t^\theta, P_t^\theta, w) \quad \forall w \in \mathbb{R}^m \quad \text{c.s. } t \in [0, T]. \quad (4.1)$$

## 4.1. Algoritmo EMSA y sus propiedades

El PMP ofrece inmediatamente un método de hallar un control  $\theta_t$  que satisfaga la condición de maximalidad (4.1). Este método, el cuál es la implementación directa de lo que establece PMP y mostrado como el Algoritmo 2, se denominará **Método de Aproximaciones Sucesivas** o **MSA** por su nombre en inglés.

---

### Algorithm 2 Algoritmo MSA

---

- 1: Inicialice control  $\theta^0$
  - 2: **for**  $k = 0$  **to**  $\#$  de iteraciones **do**
  - 3:   Resuelva  $\dot{X}_t^{\theta^k} = f(X_t^{\theta^k}, \theta_t^k) \quad X_0^{\theta^k} = X_0$
  - 4:   Resuelva  $\dot{P}_t^{\theta^k} = -\nabla_x H(X_t^{\theta^k}, P_t^{\theta^k}, \theta_t^k) \quad P_T^{\theta^k} = -\nabla \Phi(X_T^{\theta^k})$
  - 5:   Defina  $\theta_t^{k+1} = \arg \max_{w \in \mathbb{R}^m} H(X_t^{\theta^k}, P_t^{\theta^k}, w)$ .
- 

El primer paso dentro del **for**, es decir el de la línea 3, en el que se resuelve la ecuación de  $X$ , se haría hacia adelante, por ejemplo con un método de Euler. El siguiente paso igualmente pero hacia atrás, y el paso de la línea 5 posiblemente con ascenso de gradiente. Desafortunadamente, en general dicho método diverge si la inicialización  $\theta^0$  es mala [Chernousko and Lyubushin, 1982]. En parte, esto es debido a que el cambio de  $\theta^k$  a  $\theta^{k+1}$  en la línea 5 puede hacer que la dinámica de  $X^{\theta^{k+1}}, P^{\theta^{k+1}}$  sea suficientemente diferente de la de  $X^{\theta^k}, P^{\theta^k}$  de forma que  $\theta^{k+1}$  y  $\theta^{k+2}$  no sean cercanos, ni que haya convergencia de  $J(\theta^k)$ . El siguiente resultado es más específico en este respecto.

**Lema 4.1.1** ([Li et al., 2017, Lema 2]). Existe una constante  $C > 0$  tal que para cualesquiera  $\theta, \phi \in L^\infty([0, T], \Theta)$  se cumple

$$\begin{aligned} J(\phi) &\leq J(\theta) - \int_0^T \Delta_{\phi, \theta} H(t) dt \\ &\quad + C \int_0^T |f(X_t^\theta, \phi_t) - f(X_t^\theta, \theta_t)|^2 dt \\ &\quad + C \int_0^T |\nabla_x H(X_t^\theta, P_t^\theta, \phi_t) - \nabla_x H(X_t^\theta, P_t^\theta, \theta_t)|^2 dt \end{aligned}$$

donde  $\Delta_{\phi, \theta} H(t) = H(X_t^\theta, P_t^\theta, \phi_t) - H(X_t^\theta, P_t^\theta, \theta_t)$  (sólo cambia en la tercer posición, en la que se maximiza).

*Demostración.* Antes de comenzar la prueba se le recuerda al lector que en este trabajo  $C$  denota una constante, y  $k$  la constante de Lipschitz para una función que sea Lipschitz. En la mayoría de los casos se usa un subíndice para indicar el objeto que acota o función que es Lipschitz. Cuando no se indica un subíndice, la constante  $C$  o  $k$  puede cambiar de valor de línea a línea.

En general, para controles  $\phi, \theta$ , se tiene

$$J(\phi) - J(\theta) = \Phi(X_T^\phi) - \Phi(X_T^\theta) + \int_0^T L(\phi_s) - L(\theta_s) ds.$$

En virtud de

$$L(\theta_s) = P_s^\theta \dot{X}_s^\theta - H(X_s^\theta, P_s^\theta, \theta_s)$$

se tiene

$$\begin{aligned} \int_0^T L(\phi_s) - L(\theta_s) ds &= \int_0^T P_s^\theta \delta \dot{X}_s + \delta P_s \dot{X}_s^\theta + \delta P \delta \dot{X}_s ds \\ &\quad - \int_0^T H(X_s^\phi, P_s^\phi, \phi_s) - H(X_s^\theta, P_s^\theta, \theta_s) ds \end{aligned}$$

donde  $\delta X_s = X_s^\phi - X_s^\theta$ ,  $\delta P_s = P_s^\phi - P_s^\theta$ .

Por otra parte, integrando por partes

$$\begin{aligned} \int_0^T P_s^\theta \delta \dot{X}_s + \delta P_s \dot{X}_s^\theta ds &= P_s^\theta \delta X_s \Big|_{s=0}^T + \int_0^T \delta P_s \nabla_p H(X_s^\theta, P_s^\theta, \theta_s) + \delta X_s \nabla_x H(X_s^\theta, P_s^\theta, \theta_s) ds \\ &= P_s^\theta \delta X_s \Big|_{s=0}^T + \int_0^T \nabla_z H(Z_s^\theta, \theta_s) \cdot \delta Z_s ds \end{aligned}$$

donde  $Z_s^\theta = (X_s^\theta, P_s^\theta)$ . También

$$\begin{aligned} \int_0^T \delta P_s \delta \dot{X}_s ds &= \frac{1}{2} \int_0^T \delta P_s \delta \dot{X}_s ds + \frac{1}{2} \int_0^T \delta P_s \delta \dot{X}_s ds \\ &= \frac{1}{2} \delta P_s \delta X_s \Big|_0^T \\ &\quad + \frac{1}{2} \int_0^T \delta P_s \cdot \delta \nabla_p H_s + \delta X_s \cdot \delta \nabla_x H_s ds \end{aligned}$$

donde  $\delta \nabla_p H_s = \nabla_p H(Z_s^\phi, \phi_s) - \nabla_p H(Z_s^\theta, \theta_s)$ , similarmente con  $\nabla_x$ . Así,

$$\begin{aligned}
\int_0^T \delta P_s \delta \dot{X}_s ds &= \frac{1}{2} \delta P_s \delta X_s \Big|_0^T + \frac{1}{2} \int_0^T [\nabla_z H(Z_s^\phi, \phi_s) - \nabla_z H(Z_s^\theta, \theta_s)] \cdot \delta Z_s ds \\
&= \frac{1}{2} \delta P_s \delta X_s \Big|_0^T \\
&+ \frac{1}{2} \int_0^T [\nabla_z H(Z_s^\phi, \phi_s) - \nabla_z H(Z_s^\theta, \phi_s)] \cdot \delta Z_s ds \\
&+ \frac{1}{2} \int_0^T [\nabla_z H(Z_s^\theta, \phi_s) - \nabla_z H(Z_s^\theta, \theta_s)] \cdot \delta Z_s ds \\
&= \frac{1}{2} \delta P_s \delta X_s \Big|_0^T \\
&+ \frac{1}{2} \int_0^T \delta Z_s^t \nabla_z^2 H(r_1(s), \phi_s) \delta Z_s ds \\
&+ \frac{1}{2} \int_0^T [\nabla_z H(Z_s^\theta, \phi_s) - \nabla_z H(Z_s^\theta, \theta_s)] \cdot \delta Z_s ds
\end{aligned} \tag{4.2}$$

donde  $r_1(s)$  es de la forma  $Z_s^\theta + a \delta Z_s$  y se obtiene por el teorema del valor medio aplicado a

$$a \mapsto \nabla_z H(Z_s^\theta + a \delta Z_s, \phi_s) \cdot \delta Z_s$$

evaluado en  $a = 0, a = 1$ .

Por otro lado

$$\begin{aligned}
(P_s^\theta + \frac{1}{2} \delta P_s) \delta X_s \Big|_0^T &= (P_T^\theta + \frac{1}{2} \delta P_T) \delta X_T \quad \text{pues } \delta X_0 = 0 \\
&= -\nabla \Phi(X_T^\theta) \delta X_T - \frac{1}{2} [\nabla \Phi(X_T^\phi) - \nabla \Phi(X_T^\theta)] \delta X_T \\
&= -\nabla \Phi(X_T^\theta) \delta X_T - \frac{1}{2} \delta X_T^t \nabla^2 \Phi(r_2) \delta X_T \quad \text{por teorema de valor medio} \\
&= -(\Phi(X_T^\phi) - \Phi(X_T^\theta)) - \frac{1}{2} \delta X_T^t [\nabla^2 \Phi(r_2) + \nabla^2 \Phi(r_3)] \delta X_T.
\end{aligned}$$

La última igualdad surge de la expansión de Taylor con residuo para  $\Phi$  centrada en  $X_T^\theta$ . Finalmente, por Taylor

$$\begin{aligned}
H(X_s^\phi, P_s^\phi, \phi_s) - H(X_s^\theta, P_s^\theta, \theta_s) &= H(Z_s^\theta, \phi_s) - H(Z_s^\theta, \theta_s) \\
&+ \nabla_z H(Z_s^\theta, \phi_s) \delta Z_s + \frac{1}{2} \delta Z_s^t \nabla_z^2 H(r_4(s), \phi_s) \delta Z_s.
\end{aligned}$$

Así,

$$\begin{aligned}
J(\phi) - J(\theta) &= \Phi(X_T^\phi) - \Phi(X_T^\theta) + \int_0^T L(\phi_s) - L(\theta_s) ds \\
&= \Phi(X_T^\phi) - \Phi(X_T^\theta) - \int_0^T H(Z_s^\phi, \phi_s) - H(Z_s^\theta, \theta_s) ds \\
&\quad + P_s^\theta \delta X_s \Big|_0^T + \int_0^T \nabla_z H(Z_s^\theta, \theta_s) \delta Z_s ds \\
&\quad + \frac{1}{2} \delta P_s \delta X_s \Big|_0^T + \frac{1}{2} \int_0^T [\nabla_z H(Z_s^\theta, \phi_s) - \nabla_z H(Z_s^\theta, \theta_s)] \delta Z_s ds \\
&\quad \quad \quad + \frac{1}{2} \int_0^T \delta Z_s^t \nabla_z^2 H(r_1(s), \phi_s) \delta Z_s ds \\
&= -\frac{1}{2} \delta X_T^t [\nabla^2 \Phi(r_2) + \nabla^2 \Phi(r_3)] \delta X_T
\end{aligned} \tag{1}$$

$$- \int_0^T \Delta_{\phi, \theta} H(s) ds \tag{2}$$

$$+ \left(-1 + \frac{1}{2}\right) \int_0^T [\nabla_z H(Z_s^\theta, \phi_s) - \nabla_z H(Z_s^\theta, \theta_s)] \delta Z_s ds \tag{3}$$

$$+ \frac{1}{2} \int_0^T \delta Z_s^t [\nabla_z^2 H(r_1(s), \phi_s) - \nabla_z^2 H(r_4(s), \phi_s)] \delta Z_s ds. \tag{4}$$

Ahora note que para alguna constante  $C$  se cumple<sup>1</sup>

$$\begin{aligned}
|(1)| &\leq \frac{1}{2} C |\delta X_T|^2 \quad \text{pues } \nabla^2 \Phi \text{ es acotada.} \\
|(3)| &\leq \frac{1}{2} \left[ \int_0^T |\delta P_s| |f(X_s^\theta, \phi_s) - f(X_s^\theta, \theta_s)| ds \right. \\
&\quad \left. + \int_0^T |\delta X_s| |\nabla_x H(X_s^\theta, P_s^\theta, \phi_s) - \nabla_x H(X_s^\theta, P_s^\theta, \theta_s)| ds \right] \\
|(4)| &\leq \frac{C}{2} |\delta Z_s|^2.
\end{aligned}$$

En virtud de

$$\begin{aligned}
|\delta X_t| &\leq \int_0^t |f(X_s^\phi, \phi_s) - f(X_s^\theta, \theta_s)| ds \\
&\leq \int_0^t |f(X_s^\phi, \phi_s) - f(X_s^\theta, \phi_s)| + |f(X_s^\theta, \phi_s) - f(X_s^\theta, \theta_s)| ds \\
&\leq k_f \int_0^t |\delta X_s| ds + \int_0^t |f(X_s^\theta, \phi_s) - f(X_s^\theta, \theta_s)| ds \quad \text{pues } f \text{ es Lipschitz.}
\end{aligned}$$

---

<sup>1</sup>a priori la constante difiere de línea a línea. Tomando el máximo de las respectivas constantes, puede suponer sin pérdida de generalidad que es la misma.

La desigualdad de Gronwall implica que existe una constante  $C$  tal que

$$|\delta X_t| \leq C \int_0^T |f(X_s^\theta, \phi_s) - f(X_s^\theta, \theta_s)| ds.$$

Similarmente para  $\delta P_t$ ,

$$\begin{aligned} |\delta P_t| &\leq |\nabla\Phi(X_T^\phi) - \nabla\Phi(X_T^\theta)| + \int_t^T |\nabla_x H(X_s^\phi, P_s^\phi, \phi_s) - \nabla_x H(X_s^\theta, P_s^\phi, \phi_s)| ds \\ &\quad + \int_t^T |\nabla_x H(X_s^\theta, P_s^\phi, \phi_s) - \nabla_x H(X_s^\theta, P_s^\theta, \phi_s)| ds \\ &\quad + \int_t^T |\nabla_x H(X_s^\theta, P_s^\theta, \phi_s) - \nabla_x H(X_s^\theta, P_s^\theta, \theta_s)| ds \\ &\leq k_{\nabla\Phi} |\delta X_T| + \int_0^T C_p k_{D_x f} |\delta X_s| ds + \int_t^T C_{D_x f} |\delta P_s| ds \\ &\quad + \int_0^T |\nabla_x H(X_s^\theta, P_s^\theta, \phi_s) - \nabla_x H(X_s^\theta, P_s^\theta, \theta_s)| ds. \end{aligned}$$

Nuevamente por Gronwall

$$\begin{aligned} |\delta P_t| &\leq C \left[ |\delta X_T| + \int_0^T |\delta X_s| ds + \int_0^T |\nabla_x H(X_s^\theta, P_s^\theta, \phi_s) - \nabla_x H(X_s^\theta, P_s^\theta, \theta_s)| ds \right] \\ &\leq C \left[ \int_0^T |f(X_s^\theta, \phi_s) - f(X_s^\theta, \theta_s)| ds + \int_0^T |\nabla_x H(X_s^\theta, P_s^\theta, \phi_s) - \nabla_x H(X_s^\theta, P_s^\theta, \theta_s)| ds. \right] \end{aligned}$$

De esta manera, definiendo  $A = \int_0^T |f(X_s^\theta, \phi_s) - f(X_s^\theta, \theta_s)| ds$ ,  $B = \int_0^T |\nabla_x H(X_s^\theta, P_s^\theta, \phi_s) - \nabla_x H(X_s^\theta, P_s^\theta, \theta_s)| ds$ , se tiene

$$\begin{aligned} J(\phi) - J(\theta) &\leq \int_0^T -\Delta_{\phi, \theta} H(t) dt \\ &\quad + C[A^2 + [(A + B)A + AB] + (2A + B)^2] \\ &\leq \int_0^T -\Delta_{\phi, \theta} H(t) dt + C[A^2 + B^2] \end{aligned}$$

y por Cauchy Schwarz se concluye

$$\begin{aligned} J(\phi) - J(\theta) &\leq \int_0^T -\Delta_{\phi, \theta} H(t) dt \\ &\quad + \int_0^T |f(X_s^\theta, \phi_s) - f(X_s^\theta, \theta_s)|^2 ds \\ &\quad + C \int_0^T |\nabla_x H(X_s^\theta, P_s^\theta, \phi_s) - \nabla_x H(X_s^\theta, P_s^\theta, \theta_s)|^2 ds. \end{aligned}$$

Note que para esta demostración no es necesario toda la suposición **A1'**. Para la prueba sólo se usó que  $D_x f$ ,  $D_x^2 f$ ,  $\nabla\Phi$ ,  $\nabla^2\Phi$  son acotadas, y que  $D_x f$  es Lipschitz. Estas suposiciones implican que  $p$  es acotada y también las otras funciones, como  $H$ .  $\square$



El lema anterior garantiza que la función objetivo  $J$  no aumentará en el paso de optimización de MSA si el cambio en el control no modifica demasiado la dinámica de  $x$  y  $p$ .

Por otra lado, imagine el problema de maximizar el Hamiltoniano en el PMP como un problema de optimización restringido, donde las restricciones son las ecuaciones diferenciales que definen a  $x$  y  $p$ . Una idea común en dicha clase de problemas es olvidar por un momento la restricción y agregar a la función a optimizar un término de penalización por no cumplir la restricción, ver por ejemplo [Bertsekas, 1999]. Usando esta idea y el conocimiento dado por el [Li et al., 2017, Lema 2], lema 4.1.1 en este escrito, se propone usar el Hamiltoniano aumentado

$$\tilde{H}(x, p, \theta, v, q) = H(x, p, \theta) - \frac{1}{2}\rho|v - f(x, \theta)|^2 - \frac{1}{2}\rho|q + \nabla_x H(x, p, \theta)|^2$$

donde  $\rho$  es un "hiperparámetro", es decir un parámetro escogido y no obtenido por entrenamiento, para ejecutar la modificación de  $\theta^k$  a  $\theta^{k+1}$  en MSA.

El **método de aproximaciones sucesivas extendido** o **EMSA** por su nombre en inglés, es el algoritmo MSA usando  $\tilde{H}$  en vez de  $H$ , es decir

---

**Algorithm 3** Algoritmo EMSA
 

---

- 1: Inicialice control  $\theta^0$
  - 2: **for**  $k = 0$  **to**  $\#$  de iteraciones **do**
  - 3: Resuelva  $\dot{X}_t^{\theta^k} = f(X_t^{\theta^k}, \theta_t^k)$   $X_0^{\theta^k} = x_0$
  - 4: Resuelva  $\dot{P}_t^{\theta^k} = -\nabla_x H(X_t^{\theta^k}, P_t^{\theta^k}, \theta_t^k)$   $p_T^{\theta^k} = -\nabla \Phi(x_T^{\theta^k})$
  - 5: Defina  $\theta_t^{k+1} = \arg \max_{w \in \mathbb{R}^m} \tilde{H}(X_t^{\theta^k}, P_t^{\theta^k}, w, \dot{X}_t^{\theta^k}, \dot{P}_t^{\theta^k})$ .
- 

Resulta ser que usar el Hamiltoniano aumentado hace que el algoritmo EMSA converja.

**Teorema 4.1.1.** Bajo la suposición  $A1'$ , si  $\theta^0$  es un control arbitrario con  $J(\theta^0) < \infty$ , y además  $\inf_{\theta} J(\theta) > -\infty$ , entonces para  $\rho$  suficientemente grande, existe constante  $D > 0$  tal que

$$J(\theta_{\text{EMSA}}^{k+1}) - J(\theta_{\text{EMSA}}^k) \leq -D\mu_k$$

donde

$$\mu_k = \int_0^T \Delta_{\theta^{k+1}, \theta^k} H(t) dt.$$

En particular, la condición  $\inf J(\theta) > -\infty$  implica  $J(\theta^k)$  converge.

Para la discretización necesaria al implementar numéricamente EMSA es conveniente introducir la siguiente notación

$N$  = Número de pasos en la solución numérica de las ecuaciones diferenciales

$$\delta = \frac{T}{N} \quad \text{tamaño de paso}$$

$$x_n = X_{n\delta}$$

$$\vartheta_n = \theta_{n\delta}$$

$$p_n = P_{n\delta}$$

$$g(x_n, \vartheta_n) = x_n + \delta f(x_n, \vartheta)$$

$$H_d(x, p, \vartheta) = p \cdot g(x, \vartheta) - \delta L(\vartheta).$$

La forma de  $g$  y  $H_d$  es por pura conveniencia. Con esta notación las condiciones en el Principio del Máximo de Pontryagin serían

$$\begin{aligned} x_{n+1}^* &= g(x_n^*, \vartheta_n^*) & x_0^* &= x_0 \\ p_n^* &= \nabla_x H_d(x_n^*, p_{n+1}^*, \vartheta_n^*) & p_N^* &= -\nabla \Phi(x_N^*) \\ H_d(x_n^*, p_{n+1}^*, \vartheta_n^*) &\geq H_d(x_n^*, p_{n+1}^*, w) & \forall w \in \Theta, n &= 0, \dots, N-1. \end{aligned}$$

Así, la discretización del algoritmo EMSA se convierte en

---

**Algorithm 4** Algoritmo EMSA discretizado

---

- 1: Inicialice control  $\{\vartheta_n^0\}_{n=0}^{N-1}$
  - 2: **for**  $k = 0$  **to**  $\#$  de iteraciones **do**
  - 3:    $x_0^{\vartheta^k} = x_0$ .
  - 4:   **for**  $n = 0$  **to**  $N - 1$  **do**
  - 5:      $x_{n+1}^{\vartheta^k} = g(x_n^{\vartheta^k}, \vartheta_n^k)$
  - 6:      $p_N^{\vartheta^k} = -\nabla_x \Phi(x_N^{\vartheta^k})$
  - 7:     **for**  $n = N - 1$  **to**  $0$  **do**
  - 8:       $p_n^{\vartheta^k} = -\nabla_x H(x_n^{\vartheta^k}, p_{n+1}^{\vartheta^k}, \vartheta_n^k)$
  - 9:     **for**  $n = 0$  **to**  $N - 1$  **do**
  - 10:      Defina  $\vartheta_n^{k+1} = \arg \max_{w \in \Theta} H(x_n^{\vartheta^k}, p_{n+1}^{\vartheta^k}, w) - \frac{1}{2}\rho|x_{n+1}^{\vartheta^k} - g(x_n^{\vartheta^k}, w)|^2 - \frac{1}{2}\rho|p_n^{\vartheta^k} - \nabla_x H(x_n^{\vartheta^k}, p_{n+1}^{\vartheta^k}, w)|^2$ .
- 

El paso 10 en vez de hacerse duro, es decir buscar inmediatamente el máximo, se puede hacer de forma suave, por ejemplo con ascenso de gradiente. Una pregunta natural es la relación que guarda ese algoritmo con backpropagation. Al respecto se tiene la siguiente propiedad.

**Propiedad 4.1.1.** La versión a tiempo discreto del algoritmo MSA, es decir el algoritmo anterior con  $\rho = 0$ , con el paso 10 hecho por ascenso de gradiente

$$\vartheta_n^{k+1} = \vartheta_n^k + \eta \nabla_w H(x_n^{\vartheta^k}, p_{n+1}^{\vartheta^k}, w)$$

es equivalente a descenso de gradiente con backpropagation, en símbolos

$$\nabla_w H(x_n^{\vartheta^k}, p_{n+1}^{\vartheta^k}, w) = -\nabla_{\vartheta_n} J(\vartheta)$$

donde  $J(\vartheta) = \Phi(x_N^\vartheta) + \delta \sum_{i=0}^{N-1} L(\vartheta)$ .

Una consecuencia importante del resultado anterior es que para redes residuales como las que se consideran en el presente trabajo, el entrenamiento por backpropagation puede reformularse mediante el formalismo de control óptimo. En esta nueva perspectiva, la propagación de información es dirigida por  $p^\vartheta$  de una manera más simple, la ecuación diferencial. Además, el paso de maximización es por separado e independiente de cada capa. Así, en principio, éste puede paralizarse una vez calculados  $x^\vartheta, p^\vartheta$ .



# Conclusiones y Trabajo Futuro

## Conclusiones

En este manuscrito se presentaron varios aspectos en los que la comunidad matemática puede colaborar con la comunidad de aprendizaje máquina para establecer propiedades del problema de entrenamiento de redes neuronales. Para esto, primero es necesario conocer el vocabulario del área, lo cuál se presentó de manera sucinta en el primer capítulo. En el capítulo siguiente se interpretó el problema de entrenamiento como un problema de control. Para este hay dos perspectivas de estudio. Por un lado se tiene la perspectiva de Programación Dinámica, que presenta un problema rico desde la pura perspectiva matemática. Para este aún hay varios aspectos matemáticos por resolver. La otra perspectiva, Principio del Máximo de Pontryagin, permite establecer condiciones necesarias que debe cumplir un control óptimo para el problema de control. Ésta perspectiva permitió proveer de una demostración completa y detallada de un teorema sobre robustez y consistencia de los pesos obtenidos al entrenar una red residual en el capítulo 3. Finalmente, en el capítulo 4 se ejemplifica la clase de resultados en lo que respecta a algoritmos que son deseables en aprendizaje máquina, y en los que matemáticos pueden proveer un gran aporte.

Por otra parte, a raíz de lo expuesto se recaban los siguientes comentarios

- Estudiar el problema de entrenamiento de una red neuronal como la discretización de un problema de control provee de un marco teórico efectivo para establecer resultados de convergencia en probabilidad y consistencia de los parámetros encontrados al entrenar una red. No obstante, para lograr este objetivo será necesario usar la teoría de probabilidad en espacios de Banach.
- El problema 2.2, que surge de la formulación como problema de control del problema de entrenamiento de una red neuronal residual, preserva muchas similitudes y propiedades de un problema de control determinista. Esto permite obtener algunos resultados clásicos emulando las demostraciones de la situación determinista.
- La perspectiva abstracta del problema de entrenamiento como problema de control no sólo permite establecer propiedades teóricas para ambos problemas, sino que permite idear nuevos algoritmos que pueden tener muchas propiedades deseables.

## Trabajo Futuro

Para el problema teórico 2.2, hay algunos resultados pendientes por demostrar

1. Demostrar que la función de valor es la única solución de viscosidad de la ecuación de Hamilton-Jacobi-Bellman. Respecto al artículo, [E et al., 2018], hay tres pasos por hacer. Primero ver que  $v$  es Lipschitz, o al menos uniformemente continua. Luego, para probar que  $v$  es solución de viscosidad, los autores de dicho artículo se apoyan en que el Hamiltoniano es Lipschitz en cada variable. Sin embargo, el argumento ofrecido para dicho resultado tiene un serio error. El autor del presente trabajo propone intentar emular la demostración del caso determinista. Para la unicidad, es necesario una examinación muy cuidados del uso y construcción de los  $\epsilon$  usados en el artículo antes citado.
2. Demostrar que bajo ciertas condiciones de concavidad del Hamiltoniano en la tercer variable, la del control, una solución del PMP es única en un intervalo pequeño de tiempo  $[0, T_0]$ .
3. Formular y establecer una resultado de consistencia para  $J_N$  o  $\inf J_N$  respecto a  $J$  o  $\inf J$ .
4. Ver si es posible trabajar con controles en  $L^2([0, T], \mathbb{R}^m)$ , que es un espacio de Hilbert, en vez de  $L^\infty([0, T], \mathbb{R}^m)$ . Esto con el objetivo de utilizar otra versión de la desigualdad de Hoeffding, que también aparece en el artículo [Pinelis and Sakhanenko, 1986], la cuál permite retener la cuota  $\exp\left\{\frac{-Nx^2}{C_1+C_2x}\right\}$  para la convergencia de  $F_n$  y  $\theta^N$  en el teorema 3.2.2. En su defecto es establecer resultados sobre la rapidez de convergencia de  $\mathbb{E}|F_N - F|$ .

# Bibliografía

- [Ambrosio et al., 2005] Ambrosio, L., Gigli, N., and Savare, G. (2005). *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics. ETH Zürich. Birkhäuser Basel.
- [Bertsekas, 1999] Bertsekas, D. (1999). *Nonlinear programming*. Athena Scientific optimization and computation series. Athena Scientific.
- [Bressan and Piccoli, 2007] Bressan, A. and Piccoli, B. (2007). *Introduction to the Mathematical Theory of Control*, volume 2 of *AIMS Series on Applied Mathematics*. American Institute of Mathematical Sciences.
- [Cardaliaguet, 2013] Cardaliaguet, P. (2013). Notes on mean field games. Technical report, CEREMADE Université de PARIS - DAUPHINE.
- [Carmona and Delarue, 2018] Carmona, R. and Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications I*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer International Publishing, 1 edition.
- [Chernousko and Lyubushin, 1982] Chernousko, F. L. and Lyubushin, A. A. (1982). Method of successive approximations for solution of optimal control problems. *Optimal Control Applications and Methods*, 3(2):101–114.
- [E et al., 2018] E, W., Han, J., and Li, Q. (2018). A Mean-Field Optimal Control Formulation of Deep Learning. *arXiv e-prints*, page arXiv:1807.01083.
- [Fleming and Rishel, 2012] Fleming, W. and Rishel, R. (2012). *Deterministic and Stochastic Optimal Control*. Stochastic Modelling and Applied Probability. Springer New York.
- [Folland, 2013] Folland, G. (2013). *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley.
- [Goodfellow et al., 2016] Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- [Li et al., 2017] Li, Q., Chen, L., Tai, C., and E, W. (2017). Maximum Principle Based Algorithms for Deep Learning. *arXiv e-prints*, page arXiv:1710.09513.

- [Pinelis and Sakhnenko, 1986] Pinelis, I. and Sakhnenko, A. (1986). Remarks on inequalities for large deviation probabilities. *Theory of Probability & Its Applications*, 30(1):143–148.
- [Talagrand and Ledoux, 1991] Talagrand, M. and Ledoux, M. (1991). *Probability in Banach Spaces*. Classics in Mathematics. Springer-Verlag Berlin Heidelberg, 1 edition.
- [Zhang et al., 2020] Zhang, A., Lipton, Z. C., Li, M., and Smola, A. J. (2020). *Dive into Deep Learning*. <https://d2l.ai>.