UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE CIENCIAS

SECCIÓN DE POSGRADO Y SEGUNDA
ESPECIALIZACIÓN PROFESIONAL

TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAESTRO EN CIENCIAS CON MENCIÓN EN MATEMÁTICA APLICADA

TÍTULO

PRINCIPIOS DE OPTIMALIDAD DEL PROBLEMA DE CONTROL ÓPTIMO DE PROCESOS DISCRETOS

POR
TITO KAREL MANRIQUE CHUQUILLANQUI

ASESOR
DR. ELADIO OCAÑA ANAYA

LIMA-PERÚ

2011

DEDICATORIA

En memoria de mi padre, Tito,
confidente e inolvidable maestro.

AGRADECIMIENTOS

Agradezco profundamente a mis padres Tito y Ana por su gran amor, su noble ejemplo, sus consejos y por haberme brindado la oportunidad de cursar estudios superiores en Ciencias, a pesar de muchas dificultades.

También quisiera agradecer especialmente a mi asesor por su apoyo y por todas las enseñanzas que me ha dado.

RESUMEN

En la presente Tesis se estudia diversos Principios de Optimalidad para el Problema de Control Óptimo de Procesos Discretos, sistematizando las diferentes aproximaciones que lo tratan, estudiando las relaciones subyacentes entre éstas, y aportando algunos resultados nuevos en esta área.

Hemos sistematizado dentro de una visión teórica unificada las dos principales aproximaciones que estudian este problema: la Programación Dinámica, y la "discretización" del Principio del Máximo de Pontryagin.

En primer término organizamos los principales resultados de la Programación Dinámica sobre nuestro problema, logrando resumir esta aproximación en un algoritmo al que denominamos Algoritmo de Bellman.

A continuación investigamos hasta que punto la "discretización" del Principio del Máximo de Pontryagin, mantiene su validez en este nuevo contexto, para ello establecemos una distinción entre el principio discretizado débil y el fuerte. En este sentido probamos que en general el Principio del Máximo de Pontryagin sí mantiene su validez pero "débilmente". Seguidamente investigamos bajo qué condiciones mantiene su validez en el sentido "fuerte", hacemos esto desde tres perspectivas diferentes, encontrando un caso en el que la validez se mantiene expresamente (afinidad en la variable de estado).

Al realizar esta investigación, encontramos una relación directa con la Programación Dinámica, según la cual en cierto sentido el Principio de Pontryagin Discretizado es consecuencia de las Ecuaciones de Bellman.

En la última parte extendemos algunos resultados al horizonte infinito, demostrando un teorema nuevo para esta área, al que hemos denominado Principio Fuerte de Pontryagin Afín Infinito.

Contenido

Introducción

En el Problema de Control Óptimo de Procesos Discretos se busca hallar una sucesión de decisiones que controlan la evolución de un proceso, de modo tal que se maximize la suma de los beneficios obtenidos en cada decisión.

Dos son las principales aproximaciones para estudiar este problema: la Programación Dinámica, [2, 3, 5], y la "discretización" del Principio del Máximo de Pontryagin, [7, 8, 6], el cual es un teorema central en la Teoría de Control de Procesos Continuos [20].

Debido a que la Programación Dinámica esta desarrollada en gran cantidad de libros, [2, 3, 5, 14, 21, 23] en esta Tesis sólo resumimos y sistematizamos los hechos principales que nos atañen, para luego concentrarnos sobre todo en investigar exhaustivamente hasta donde la "discretización" del Principio del Máximo de Pontryagin, nos es útil para estudiar las soluciones de nuestro problema, tema que anteriormente ha sido investigado poco.

En esta investigación hemos encontrado que aunque dicho principio "discretizado" ya no es válido para los problemas discretos (en general), continua siendo cierto, de un modo "debilitado", en contextos bastante generales.

Además hemos estudiado con detenimiento bajo qué condiciones dicho principio sigue siendo cierto sin necesidad de "debilitarlo", y también en que casos llega a ser incluso una condición suficiente. Al realizar esta investigación hemos encontrado una interesante interrelación entre la Programación Dinámica y esta aproximación, lo que nos da otra perspectiva para interpretar esta última.

La gran importancia que tiene en la Teoría de Control de Procesos Continuos, el Principio del Máximo de Pontryagin, [8, 20], motivó la realización de todas estas investigaciones. Buscábamos comprender y establecer de un modo más explícito la relación intuitiva entre este principio y su versión "discretizada". Consideramos que este objetivo ha sido conseguido, aún cuando hay todavía mucho terreno por explorar.

A continuación describiremos cada uno de los capítulos de esta Tesis.

En el Capítulo 1 establecemos la definición precisa del Problema de Control Óptimo de Procesos Discretos (en Horizonte Finito e Infinito) y las notaciones que usaremos en el resto de la Tesis. También describimos las diferentes aproximaciones que nos permiten resolverlo y en la última parte demostramos algunos teoremas generales de existencia de soluciones óptimas.

En el Capítulo 2 sintetizamos todos los resultados más importantes de la Programación Dinámica que se usan para resolver el Problema de Control Óptimo Discreto en Horizonte Finito, logrando resumir toda la metodología de solución en un algoritmo al que denominamos "Algoritmo de Bellman".

En el Capítulo 3 "discretizamos" el Principio del Máximo de Pontryagin obteniendo una versión discreta a la que denominamos Principio Fuerte de Pontryagin Discreto (PFP), a su vez, debilitándo a éste, definimos los Principios Débiles de Pontryagin (PdP). A lo largo de este capítulo investigamos contextos diversos en los que estos últimos son válidos. Concluímos este capítulo probando, con un contraejemplo, que en general el PFP no es cierto.

En el Capítulo 4 estudiamos con detenimiento bajo que condiciones el PFP es una condición necesaria de óptimo. Encontramos que existen diferentes contextos en los que lo es, pero de un modo aproximado. Entre todos estos contextos resalta el caso en que el problema es "afín respecto a la variable de estado" pues en este caso la válidez es expresa y no sólo aproximada.

Investigando con más detenimiento cuando es válido el PFP hemos encontrado una interrelación entre la Programación Dinámica y el PFP, mas precisamente entre las Ecuaciones de Bellman y subproblemas de optimización del PFP, todo esto nos proporciona una perspectiva alternativa para interpretar y comprender el PFP.

Por último en el Capítulo 5 extendemos algunos de los resultados anteriores para estudiar problemas de Control Óptimo de Procesos Discretos de Horizonte Infinito, consiguiendo generalizar bajo ciertas hipótesis diversos principios débiles del Capítulo 3. Terminamos este capítulo demostrando que el PFP es válido también en el contexto infinito bajo ciertas condiciones de afinidad en la variable de estado. Este resultado nos parece bastante interesante.

Finalmente digamos que una razón importante que nos motivó a realizar esta investigación es la gran ausencia de libros que traten de un modo sistemático y específico el Problema de Control de Procesos Discretos, hecho sorprendente debido a la gran cantidad de problemas reales que se pueden modelar dentro de este contexto. La presente Tesis busca subsanar de alguna manera esta situación y ser base para estudios posteriores en los que por ejemplo se aborde el estudio de la Teoría de Control Óptimo de Procesos Discretos Estocásticos, sobre la cual tampoco no hay muchos libros que la traten específicamente.

Capítulo 1

Control Óptimo de Procesos Discretos

En el presente Capítulo definiremos el problema de Control Óptimo de Procesos Discretos, tanto para horizonte finito, como para infinito, así como los problemas relacionados.

Describirémos también las diferentes formas de abordar este problema, para hallar las soluciones óptimas. Y finalmente demostraremos algunos teoremas generales de existencia de solución.

1.1 Planteamiento del Problema

Comencemos nuestro trabajo presentando el problema de control óptimo de procesos discretos y sus variantes.

Una primera aproximación intuitiva al problema es considerar el siguiente ejemplo. Supongamos que tenemos un objeto "x" cuya posición es descrita por n coordenadas ($x \in \mathbb{R}^n$). En el tiempo inicial ($t = 0$) el objeto se encuentra en la posición $\hat{x}_0$. De otro lado, tenemos ciertos mecanismos (controles) con los cuales podemos desplazar el objeto. También tenemos funciones definidas para los diferentes instántes de decisión que nos dan una valoración (premios o beneficios) según el control usado y la posición en que se aplicaron. Similarmente una última función nos valora el estado final del objeto.

Añadamos además la restricción de que sólo tenemos "T" decisiones que podemos ejecutar ($T \in \mathbb{N}$, problema finito). Surge entonces un problema de optimización natural: Dado un punto inicial $\hat{x}_0$, hallar una sucesión de decisiones $(u_0, \cdots, u_{T-1})$, de modo que el beneficio total (suma de los beneficios instantáneos y beneficio final) sea el máximo.

Matemáticamente el problema es formulado de un modo más general

$$\begin{array}{ll} \max & B_T(x_0,\cdots,x_T,u_0,\cdots,u_{T-1}) := \sum_{t=0}^{T-1} f_t^0(x_t,u_t) + f_T^0(x_T) \\ \text{s.a.} & \text{para todo } \ t=0,\cdots,T-1, \\ & x_{t+1}\in X_{t+1}, \ \ u_t\in U_t, \\ & x_{t+1}=f_t(x_t,u_t), \ x_0=\hat{x}_0\in X_0, \end{array} \qquad (\mathcal{P}_T(\hat{x}_0))$$

donde para $t=0,\cdots,T-1$,

- $X_t \subseteq \mathbb{R}^n$ y $X_T \subseteq \mathbb{R}^n$ (posiciones admisibles del objeto en el instante t),
- $U_t \subseteq \mathbb{R}^m$ (los controles a usarse en el instante t),
- $f_t^0 : X_t \times U_t \to \mathbb{R}$ (el beneficio de las decisiones en cada posición e instante),
- $f_t : X_t \times U_t \to \mathbb{R}^n$ (la dinámica de cómo actua el control sobre el objeto),
- $f_T^0 : X_T \to \mathbb{R}$ (el beneficio de la posición final).

El conjunto de los $(x_0,\cdots,x_T,u_0,\cdots,u_{T-1})$ satisfaciendo el sistema en el problema $\mathcal{P}_T(\hat{x}_0)$ es llamado conjunto admisible y es denotado por $\mathrm{Adm}_T(\hat{x}_0)$,

$$\left\{ (x_0,\cdots,x_T,u_0,\cdots,u_{T-1}) \;\middle|\; \begin{array}{l} \text{para todo } \ t=0,\cdots,T-1, \\ x_{t+1}\in X_{t+1}, \ \ u_t\in U_t, \\ x_{t+1}=f_t(x_t,u_t) \ \text{y} \ x_0=\hat{x}_0\in X_0 \end{array} \right\}.$$

A este problema le denominaremos “Problema de Control Óptimo Discreto en Horizonte Finito”.

Una observación muy importante que nos servirá mucho más adelante es que dada una sucesión de decisiones $u_0,\cdots,u_{T-1}$, se genera una única sucesión de estados $x_1,\cdots,x_T$, a partir de $\hat{x}_0$. A esta sucesión la llamaremos “proceso”.

Evidentemente nuestra formulación es mucho más general que el problema del ejemplo anterior, pues consideramos aquí entre otras cosas que los “estados” (posiciones) del objeto están limitados a ciertos conjuntos del espacio, que cambian con el tiempo. Similar condición pedimos para los controles.

De otro lado, el mecanismo con el que actua el control sobre el objeto cambia con el tiempo (por ejemplo, a causa de un desgaste, etc.). También tenemos distintas formas de medir la utilidad según el instánte.

Ahora bien, la formulación presentada puede ser interpretada en un contexto muchísimo más amplio. Para ello en vez de que “x” represente las coordenadas espaciales de un “objeto”, podemos interpretarlo como las coordenadas generalizadas (terminología usual en mecánica clásica) de un estado de un sistema dinámico (o incluso de un sistema en una caracterización más abstracta). De otro lado los "controles" serían decisiones que se toman para alterar el sistema de modo que obtengamos ciertos "beneficios" de estas decisiones.

Así pues, la formulación dada modela de un modo general el problema de maximizar el beneficio obtenido al realizar T decisiones que alteran la evolución de un sistema. Es claro que intuitivamente este problema generaliza el problema de optimización estática el cual puede considerarse como un problema de control de una sola decisión.

En este marco general podemos estudiar por ejemplo el problema del crecimiento óptimo de la economía de una nación usando modelos simplificados de la evolución de las variables macroeconómicas a lo largo de los años. Las diferentes políticas económicas se entenderían como los controles y, los beneficios, serían las valoraciones del crecimiento económico (por ejemplo PBI, etc.), estos problemas son estudiados en la Macreconomía Dinámica [18, 22, 21].

Existen muchos ejemplos de posibles aplicaciones de esta formulación, básicamente se prodría tomar toda sucesión de T decisiones que busque optimizar la suma de beneficios obtenidos por cada decisión siempre que se conozca con precisión (problema determinista) cómo afecta el control al estado y las funciones beneficio.

Señalemos más generalmente algunas variaciones de la formulación anterior. En primer lugar observemos que la posición final cumple $x_T \in X_T$. Así pues estamos dando una restricción sobre el paradero final del sistema. Si el conjunto X_T es unitario, entonces se dirá que el problema es con **extremo derecho fijo** y se le denotará como $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$. Contraponiendose a los otros casos, X_T no unitario, a los que se les llama de **extremo derecho libre**. Al conjunto X_T se le suele llamar "conjunto de objetivos".

Por otra parte, tenemos que los conjuntos de controles U_t no dependen de x_t. No obstante, en otras situaciones si hay tal dependencia, es decir $U_t(x_t)$. Es claro que este caso es más general. Este planteamiento es más natural (las decisiones que podamos tomar dependen de donde nos encontramos), pero es más difícil de tratar.

Finalmente, hasta aquí hemos tratado el problema de control para T decisiones. Evidentemente se ha supuesto que $T < \infty$, por ello se conoce a estos problemas como de **Horizonte Finito**. Cuando $T = \infty$, el problema se llama de **Horizonte Infinito**. Los problemas de horizonte infinito surgen por ejemplo en la gestión de recursos renovables limitados [9, 11, 12]. En efecto, si este no fuese así, es decir si estos fuesen considerados como problemas de horizonte finito, estaríamos asumiendo que sólo nos importa obtener beneficios hasta un tiempo finito, **olvidándonos de las generaciones futuras**, que en principio las hemos de suponer y "desear" ilimitadas en el tiempo. Como observación tangente digamos que precisamente éste es uno de los problemas más importantes sobre política económica en la actualidad, pues la extinción de miles de especies y la depredación del planeta tierra hace urgente estudiar la viabilidad de lo que se ha llamado el **Desarrollo Sostenible**, es decir un desarrollo que sea posible ad infinitum, por lo cual es un problema de infinitas decisiones que cada generación irá tomando, sin terminar los recursos, pensando en las futuras que vendrán. Estas reflexiones fundamentan la necesidad de plantear el siguiente problema:

$$\begin{array}{rl} \max & B((x_0, x_1, \cdots), (u_0, \cdots)) := \sum_{t=0}^{\infty} f_t^0(x_t, u_t) \\ \text{s.a} & \text{para todo} \quad t \geq 0, \; x_t \in X_t, \; u_t \in U_t, \\ & x_{t+1} = f_t(x_t, u_t), \; x_0 = \hat{x}_0 \in X_0. \end{array} \qquad (\mathcal{P}(\hat{x}_0))$$

donde, similar al caso de horizonte finito, para todo $t \geq 0$,

- $X_t \subseteq \mathbb{R}^n$, $U_t \subseteq \mathbb{R}^m$,
- $f_t^0 : X_t \times U_t \to \mathbb{R}, \; f_t : X_t \times U_t \to \mathbb{R}^n$.

El conjunto de los $((x_0, x_1, \cdots), (u_0, \cdots))$ satisfaciendo el sistema en el problema $(\mathcal{P}(\hat{x}_0))$ es llamado conjunto admisible y es denotado por $\operatorname{Adm}(\hat{x}_0)$,

$$\operatorname{Adm}(\hat{x}_0) := \left\{ ((x_0, x_1, \cdots), (u_0, \cdots)) \;\middle|\; \begin{array}{l} \text{para todo } \; t \geq 0, \, x_t \in X_t, \; u_t \in U_t, \\ x_{t+1} = f_t(x_t, u_t) \; \text{ y } \; x_0 = \hat{x}_0 \in X_0. \end{array} \right\}.$$

Observación. En el planteamiento anterior hemos usado el criterio

$$\max \sum_{t=0}^{\infty} f_t^0(x_t, u_t),$$

es decir, definimos como sucesiones óptimas aquellas que maximizen la sumatoria anterior. Otros dos criterios de optimalidad que también serán vistos en esta Tesis son los siguientes: se dice que $((\bar{x}_0, \bar{x}_1, \cdots), (\bar{u}_0, \cdots)) \in \operatorname{Adm}(\hat{x}_0)$ es óptimo si para todo $((x_0, x_1, \cdots), (u_1, \cdots)) \in \operatorname{Adm}(\hat{x}_0)$, se cumple:

$$(\mathcal{L}, \mathcal{I}) \qquad \liminf_{T \to \infty} \left(\sum_{t=0}^{T} f_t^0(\bar{x}_t, \bar{u}_t) - \sum_{t=0}^{T} f_t^0(x_t, u_t) \right) \geq 0,$$

o también,

$$(\mathcal{L}, \mathcal{II}) \qquad \limsup_{T \to \infty} \left(\sum_{t=0}^{T} f_t^0(\bar{x}_t, \bar{u}_t) - \sum_{t=0}^{T} f_t^0(x_t, u_t) \right) \geq 0.$$

Según usemos cada uno de estos critérios (para definir una sucesión óptima) podremos definir los problemas $\mathcal{P}_{\mathcal{I}}(\hat{x}_0)$ para $(\mathcal{L}, \mathcal{I})$ y $\mathcal{P}_{\mathcal{II}}(\hat{x}_0)$ para $(\mathcal{L}, \mathcal{II})$.

Problema de control continuo con horizonte finito: Aún cuando esta Tesis trata el caso discreto, consideramos necesario al menos señalar algunos elementos del problema continuo. Históricamente este fue el caso más estudiado y su planteamiento es anterior al problema discreto. Por eso, como veremos más adelante, se toma el caso continuo y los métodos que se usan en este, como una fuente de ideas heurísticas para la solución del problema discreto.

De un modo similar al caso discreto, el problema continuo que presentaremos se puede interpretar como el de maximizar un cierto criterio respecto a las diferentes evoluciones que tiene un sistema dinámico al elegir diversos controles (caminos) en un intervalo de tiempo $[0, T]$ comenzando a partir de un punto inicial $\hat{x}_0$.

De un modo algo informal presentemos el problema como sigue, se supondran que las funciones ahí mencionadas cumplen con las condiciones de regularidad de modo que el problema esté bien definido.

Sean $T > 0$, $\hat{x}_0 \in \Omega \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$, $f : \Omega \times U \to \mathbb{R}^n$, $f^0 : \Omega \times U \to \mathbb{R}$ y $f_T^0 : \Omega \to \mathbb{R}$. El problema de control continuo (con horizonte T) puede ser formulado como sigue:

$$\begin{array}{ll} \max & B(x(\cdot), u(\cdot)) := \int_0^T f^0(x(t), u(t))dt + f_T^0(x_T), \\ \text{s.a.} & \qquad (x(\cdot), u(\cdot)) \in \text{Adm}\,(\hat{x}_0), \end{array} \qquad (\mathcal{PC}_T(\hat{x}_0))$$

donde,

$$\text{Adm}\,(\hat{x}_0) := \left\{ (x(\cdot), u(\cdot)) \;\middle|\; \begin{array}{l} u : [0, T] \to U,\, x : [0, T] \to \Omega,\ \forall\, t \in [0, T], \\ \dot{x}(t) = f(x(t), u(t)),\ x(0) = \hat{x}_0,\ x(T) \in M_f \end{array} \right\},$$

con $M_f \subseteq \mathbb{R}^n$.

En este esquema continuo trabajaremos en el intervalo $0 \leq t \leq T$ y se denotarán con "x" a los elementos de Ω.

1.2 Metodología de Solución

Existen dos métodos principales para tratar de encontrar la solución de un problema de control óptimo discreto (o al menos aproximarnos a ella). Ambos se basan en caracterizar a la solución óptima mediante alguna propiedad que ésta posee, y de esta manera poder descartar aquellos procesos que no satisfacen dicha propiedad. Este nos permite acercarnos más a la posible solución, o incluso llegar a ella.

1) El primer método se basa en usar el Principio de Optimalidad de Bellman, el cual dice: "**dada una sucesión óptima de decisiones, toda**

subsucesión de ésta es, a su vez, óptima". El sentido exacto de esta afirmación se explicará en el capítulo siguiente.

Basándonos en este principio podemos establecer dos caminos para abordar la busqueda de la solución del problema: a) podemos definir la metodología de la Programación Dinámica considerando subsucesiones iniciales o terminales de un modo recursivo y haciendo uso a la vez de las funciones valor (ver Capítulo 2). Así, de esta manera, se obtiene una caracterización de las soluciones óptimas e incluso un algoritmo para hallarlas. b) podemos establecer las ecuaciones de "cálculo variacional discreto" concentrándonos en las subsucesiones intermedias y así por lo tanto, obtener una propiedad que caracterize a las soluciones óptimas análoga a las ecuaciones de Euler- Lagrange para el cálculo variacional continuo Sin embargo esta aproximación es en esencia un caso particular del segundo método.

2) El segundo método consiste en ver al proceso entero de decisiones como un punto en un espacio más grande, el espacio de procesos, y usar en este caso los teoremas de condiciones necesarias y suficientes de optimalidad (Programación no Lineal, etc.). Por lo cual podemos obtener condiciones necesarias y suficientes para los procesos óptimos. La particularidad con la que está definido el problema de control permitirá interpretar estas condiciones de una manera especial que nos induce a observar estos resultados como análogos al principio del máximo de Pontryagin para procesos continuos (Capítulo 3). Un estudio más profundo nos mostrará que de hecho este principio se cumple bajo ciertos condiciones, más aún en algunos casos este método es el más eficiente para hallar la solución.

Antes de continuar con los diferentes métodos, demostremos un teorema de existencia de solución, que nos servirá para sustentar nuestra investigación (una vez que se sabe que existe una solución podemos pasar a intentar buscarla). La metodología para demostrar este resultado nos será útil en los capítulos posteriores (Capítulos 3 y 4).

Teorema 1.2.1 (Teorema de Existencia) *Para el problema $P_T(\hat{x}_0)$ asumamos las hipótesis siguientes:*

- *Para $t = 0, \cdots, T-1$, los conjuntos U_t y X_{t+1} son compactos.*
- *Para $t = 0, \cdots, T-1$, las funciones f_t^0, f_t y f_T^0 son continuas en sus respectivos dominios.*
- $\mathrm{Adm}_T(\hat{x}_0) \neq \emptyset$.

Entonces el problema $P_T(\hat{x}_0)$ admite una solución.

Prueba. La idea general de la prueba consiste en transformar el problema dinámico en uno estático con su función objetivo continua y su conjunto factible compacto, de modo tal que podemos aplicar el Teorema de Weierstrass que asegura la existencia de una solución.

Con respecto al conjunto admisible: Denotemos $Z = \prod_{t=1}^{T} X_t \times \prod_{t=0}^{T-1} U_t$, entonces

$$\operatorname{Adm}_T(\hat{x}_0) = \left\{ (x_0, \cdots, x_T, u_0, \cdots, u_{T-1}) \in Z \;\middle|\; \begin{array}{l} x_1 = f_0(\hat{x}_0, u_0), \\ \forall\ t = 1, \cdots, T-1, \\ x_{t+1} = f_t(x_t, u_t). \end{array} \right\}.$$

Consideremos la función $F : Z \to (\mathbb{R}^n)^T$ definida por

$$F((x_1, \cdots, x_T, u_0, \cdots, u_{T-1})) = (x_1 - f_0(\hat{x}_0, u_0), \cdots, x_T - f_{T-1}(x_t, u_T)).$$

Por hipótesis, F es continua y, debido a que $\operatorname{Adm}_T(\hat{x}_0) = F^{-1}(\{0\})$, el conjunto admisible $\operatorname{Adm}_T(\hat{x}_0)$ es cerrado. De otro lado, por las hipótesis de compacidad, Z es compacto y, debido a que $\operatorname{Adm}_T(\hat{x}_0) \subseteq Z$, entonces $\operatorname{Adm}_T(\hat{x}_0)$ es compacto.

Con respecto a la función objetivo: Definamos la función $F^0 : Z \to \mathbb{R}$ por

$$F^0(x_1, \cdots, x_T, u_0, \cdots, u_{T-1}) = \sum_{t=0}^{T-1} f_t^0(x_t, u_t) + f_T^0(x_T).$$

Por hipótesis, F^0 es continuo.

Por lo tanto, por el teorema de Weierstrass, existe una solución del problema

$$\max_{z \in \operatorname{Adm}_T(\hat{x}_0)} F^0(z)$$

que coincide con el problema $P_T(\hat{x}_0)$. ■

Teorema 1.2.2 (2do Teorema de Existencia) *Para el problema $P_T(\hat{x}_0)$ asumamos las hipótesis siguientes:*

- *Para $t = 0, \cdots, T-1$, los conjuntos U_t son compactos.*
- *Para $t = 0, \cdots, T-1$, se cumple $f_t(X_t \times U_t) \subseteq X_{t+1}$.*
- *Para $t = 0, \cdots, T-1$, las funciones f_t^0, f_t y f_T^0 son continuas en sus respectivos dominios.*
- *$Adm_T(\hat{x}_0) \neq \emptyset$.*

Entonces el problema $P_T(\hat{x}_0)$ admite una solución.

Prueba. Teniendo encuenta la demostración anterior, el teorema se reduce a mostrar que $\operatorname{Adm}_T(\hat{x}_0)$ es compacto.

Denotemos

$$K_1 = f_0(\hat{x}_0, U_0),$$

entonces K_1 es compacto y $K_1 \subset X_1$. Para $t = 1, \cdots, T-1$, denotemos

$$K_{t+1} = f_t(X_t \times K_t),$$

luego K_{t+1} es compacto y por lo tanto $\prod_{t=1}^{T} K_t \times \prod_{t=0}^{T-1} U_t$ también lo es y contiene a $\operatorname{Adm}_T(\hat{x}_0)$. Por el argumento de la demostración anterior, $\operatorname{Adm}_T(\hat{x}_0)$ es cerrado. Se sigue entonces que $\operatorname{Adm}_T(\hat{x}_0)$ es compacto. ■

Un caso al que puede aplicarse el teorema anterior es cuando para $t = 0, \cdots, T$, $X_t = \mathbb{R}^n$ y, para $t = 0, \cdots, T-1$, los conjuntos U_t son compactos y las funciones f_t^0, f_t y f_T^0 son continuas. Este caso es bastante usado.

Los resultados anteriores nos aseguran la existencia de solución bajo ciertas condiciones. El siguiente paso será buscar algunos métodos para llegar a la solución del problema. En el capítulo siguiente estudiaremos la Programación Dinámica, con el cual se logra resolver los problemas de control óptimo de procesos discretos en horizonte finito.

Capítulo 2

Programación Dinámica

En este capítulo sintetizamos todos los resultados de la Programación Dinámica [2, 18, 23] que son usados para resolver problemas de Control Óptimo de Procesos Discretos en Horizonte Finito.

Logramos resumir toda la metodología de la Programación Dinámica para estos problemas en un algoritmo al que denominamos "Algoritmo de Bellman".

2.1 Principio de Optimalidad de Bellman

Recordemos nuestro problema de control $\mathcal{P}_T(\hat{x}_0)$, cuya representación simplificada es

$$\max_{z \in \text{Adm}_T(\hat{x}_0)} B_T(z),$$

donde $\text{Adm}_T(\hat{x}_0)$ es el conjunto

$$\left\{ z := (x_0, \cdots, x_T, u_0, \cdots, u_{T-1}) \;\middle|\; \begin{array}{l} \text{para todo } \; t = 0, \cdots, T-1, \\ x_{t+1} \in X_{t+1}, \quad u_t \in U_t, \\ x_{t+1} = f_t(x_t, u_t) \;\text{ y }\; x_0 = \hat{x}_0 \in X_0 \end{array} \right\}.$$

La Programación Dinámica surge como un método de resolución de estos problemas. Fue propuesta y desarrollada sobre todo por Richard Bellman, quien de la observación de una propiedad particular de los procesos óptimos y su relación con las funciones que él denominó de "valor", obtuvo una serie de ecuaciones a partir de las cuales se puede caracterizar la solución para el problema.

Comenzaremos con esta propiedad especial, que hoy en día es conocida con el nombre de **Principio de Optimalidad de Bellman**, que dice: **dada una**

sucesión óptima de decisiones, toda subsucesión de ésta es, a su vez, óptima.

Antes de continuar el estudio de la programación dinámica señalemos un resultado que simplificará grandemente las notaciones.

Lema 2.1.1 *Dado un elemento* $(x_0, \cdots, x_T, u_0, \cdots, u_{T-1}) \in \mathrm{Adm}_T(\hat{x}_0)$ *entonces* $(x_1, \cdots, x_T)$ *es unívocamente determinado por* $(u_0, \cdots, u_{T-1})$.

Prueba. De su definición con u_0 se determiná unívocamente $x_1 = f_0(\hat{x}_0, u_0)$. A su vez con x_1 y u_1 se determina un único $x_2 = f_1(x_1, u_1)$, procediendo sucesivamente, encontramos una única sucesión $(x_1, \cdots, x_T)$ generada por las decisiones consecutivas $(u_0, \cdots, u_{T-1})$. ■

Teniendo en cuenta el lema anterior vamos a redefinir el conjunto admisible.

$$\mathrm{Adm}_T(\hat{x}_0) := \left\{ (u_0, \cdots, u_{T-1}) \;\middle|\; \begin{array}{l} \text{existe } (x_0, \cdots, x_T) \text{ tal que} \\ (x_0, \cdots, x_T, u_0, \cdots, u_{T-1}) \in \mathrm{Adm}_T(\hat{x}_0) \\ \text{(de la definición precedente)} \end{array} \right\}.$$

Así, la función objetivo B_T será reemplazada de manera equivalente por

$$B_T(\hat{x}_0; u_0, \cdots, u_{T-1}) := B_T(\hat{x}_0, \cdots, x_T, u_0, \cdots, u_{T-1}).$$

En general cuando no se diga lo contrario, siempre se supondra que los decisiones u_t determinan las posiciones x_t. También algunas veces diremos que la sucesión de decisiones $(u_1, \cdots, u_T)$ es solución del problema $\mathcal{P}_T(\hat{x}_0)$. Notemos que no se puede proceder recíprocamente, es decir, dado $(x_0, \cdots, x_T)$ puede existir varias sucesiones de decisiones que lo originan.

Pasemos inmediatamente a describir en un sentido concreto (matemático) el principio de optimalidad de Bellman.

Asociado al problema $\mathcal{P}_T(\hat{x}_0)$ definimos, para cada $k = 0, \cdots, T-1$, los subproblemas,

$$\max_{w \in \mathrm{Adm}_T^k(\hat{x}_k)} B_T^k(w) := \sum_{t=k}^{T-1} f_t^0(x_t, u_t) + f_T^0(x_T), \qquad (\mathcal{P}_T^k(\hat{x}_k))$$

donde

$$\mathrm{Adm}_T^k(\hat{x}_k) := \left\{ w = (u_k, \cdots, u_{T-1}) \;\middle|\; \begin{array}{l} \text{para todo } t = k, \cdots, T-1, \\ x_{t+1} \in X_{t+1}, \quad u_t \in U_t, \\ x_{t+1} = f_t(x_t, u_t) \text{ y } x_k = \hat{x}_k \in X_k \end{array} \right\}.$$

Por definición, el problema $\mathcal{P}_T^0(\hat{x}_0)$ coincide con $\mathcal{P}_T(\hat{x}_0)$ y también $B_T^0 = B_T$.

Lema 2.1.2 (Principio de Optimalidad de Bellman) *Asumamos que la sucesión* $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ *es una solución óptima del problema* $\mathcal{P}_T(\hat{x}_0)$ *con proceso asociado* $(\hat{x}_0, \cdots, \hat{x}_T)$*. Entonces, para* $k = 0, \cdots, T-1$, $(\hat{u}_k, \cdots, \hat{u}_{T-1})$*, es también una solución óptima del subproblema* $\mathcal{P}_T^k(\hat{x}_k)$*.*

Prueba. Por definición, $(\hat{u}_k, \cdots, \hat{u}_{T-1}) \in \mathrm{Adm}\,_T^k(\hat{x}_k)$. Supongamos por contradicción que $(\hat{u}_k, \cdots, \hat{u}_{T-1})$ no es óptima de $\mathcal{P}_T^k(\hat{x}_k)$. Entonces existe $(\tilde{u}_k, \cdots, \tilde{u}_{T-1}) \in \mathrm{Adm}\,_T^k(\hat{x}_k)$ tal que

$$B_T^k(\hat{x}_k; \tilde{u}_k, \cdots, \tilde{u}_{T-1}) > B_T^k(\hat{x}_k; \hat{u}_k, \cdots, \hat{u}_{T-1}).$$

Definamos el proceso $(\tilde{x}_0, \cdots, \tilde{x}_T)$ por,

$$\begin{aligned} \tilde{x}_0 &= \hat{x}_0, \\ \tilde{x}_{t+1} &= f_t(\tilde{x}_t, \hat{u}_t), \quad t = 0, \cdots, k-1, \\ \tilde{x}_{t+1} &= f_t(\tilde{x}_t, \tilde{u}_t), \quad t = k, \cdots, T-1. \end{aligned}$$

De esto claramente se tiene que $(\hat{u}_0, \cdots, \hat{u}_{k-1}, \tilde{u}_k, \cdots, \tilde{u}_{T-1}) \in \mathrm{Adm}\,_T(\hat{x}_0)$. Se cumple

$$\begin{aligned} B_T(\hat{x}_0; \hat{u}_0, \cdots, \hat{u}_{k-1}, \tilde{u}_k, \cdots, \tilde{u}_{T-1}) &= \textstyle\sum_{t=0}^{k-1} f_t^0(\hat{x}_t, \hat{u}_t) + B_T^k(\hat{x}_k; \tilde{u}_k, \cdots, \tilde{u}_{T-1}) \\ &> \textstyle\sum_{t=0}^{k-1} f_t^0(x_t, u_t) + B_T^k(\hat{x}_k; \hat{u}_{k+1}, \cdots, \hat{u}_T) \\ &= B_T(\hat{x}_0; \hat{u}_0, \cdots, \hat{u}_T). \end{aligned}$$

Este contradice la optimalidad de $(\hat{u}_0, \cdots, \hat{u}_T)$. Se deduce que $(\hat{u}_k, \cdots, \hat{u}_T)$ es una solución óptima de $\mathcal{P}_T^k(\hat{x}_k)$. ■

2.2 Ecuaciones de Bellman

Para cada $k = 1, \cdots T-1$, asociada al problema $\mathcal{P}_T^k(x_k)$, definamos la función $V_k : X_k \to \mathbb{R}$,

$$V_k(x) := \sup_{\mathrm{Adm}\,_T^k(x)} B_T^k(x; u_k, \cdots, u_{T-1}).$$

La función V_k es llamada la **función Valor** asociada al problema $\mathcal{P}_T^k(x_k)$.

Observemos que las funciones V_k maximizan el valor a partir de las decisiones k-ésima en adelante. Por ello en la etapa T, al no tener decisiones posteriores (de $T+1$, en adelante), le asigamos por convención $V_T(\cdot) = f_T^0(\cdot)$.

Con el principio de optimalidad de Bellman, probaremos que estas funciones satisfacen un sistema de "ecuaciones" al que llamaremos **Sistema de Bellman**. Demostraremos que las "soluciones" de este sistema caracterizan las soluciones óptimas obteniendo una condición necesaria y suficiente de optimalidad.

Proposición 2.2.1 (Ecuaciones de Bellman) *Para $k = 0, \cdots, T-1$, las funciones V_k satisfacen*

$$V_k(x) = \sup_{u \in U_k} \{f_k^0(x,u) + V_{k+1}(f_k(x,u))\}, \qquad (EB_k)$$

con $V_T(x) = f_T^0(x)$ para todo x. La ecuación (EB_k) es llamada una **ecuación de Bellman**.

Prueba. Tomemos $k \in \{0, \cdots, T-1\}$ y $x \in X_k$ fijo y arbitrario. Demostremos que se cumple la igualdad en dos partes. Primero dado $u \in U_k$ fijo arbitrario, por la definición de V_{k+1} como un supremo, dado $\epsilon > 0$ existe una sucesión de decisiones $(\hat{u}_{k+1}, \hat{u}_{k+2}, \cdots, \hat{u}_{T-1})$ tal que

$$V_{k+1}(f_k(x,u)) - \epsilon < B_T^{k+1}(f_k(x,u); \hat{u}_{k+1}, \hat{u}_{k+2}, \cdots, \hat{u}_{T-1}).$$

Sumando a ambos lados $f_k^0(x,u)$ deducimos que $(\hat{u}_{k+1}, \cdots, \hat{u}_{T-1})$ satisface

$$f_k^0(x,u) + V_{k+1}(f_k(x,u) - \epsilon < f_k^0(x,u) + B_T^{k+1}(f_k(x,u); \hat{u}_{k+1}, \cdots, \hat{u}_{T-1}) \leq V_k(x).$$

Como $\epsilon > 0$ puede ser tomado arbitrariamente pequeño. Deducimos que para cualquier $u \in U_k$ se cumple

$$f_k^0(x,u) + V_{k+1}(f_k(x,u)) \leq V_k(x),$$

por lo tanto

$$\sup_{u \in U_k} \left[f_k^0(x,u) + V_{k+1}(f_k(x,u))\right] \leq V_k(x).$$

Para la desigualdad contraria tomamos nuevamente $\epsilon > 0$ arbitrario. Desde que V_k es un supremo, existe una sucesión de decisiones $(\hat{u}_k, \cdots, \hat{u}_{T-1})$ tal que

$$V_k(x) - \epsilon \leq B_T^k(x; \hat{u}_k, \cdots, \hat{u}_{T-1}).$$

Por la definición de B_T^{k+1} es claro que

$$\begin{aligned} B_T^k(x; \hat{u}_k, \cdots, \hat{u}_{T-1}) &= f_k^0(x, \hat{u}_k) + B_T^{k+1}(f_k(x, \hat{u}_k); \hat{u}_{k+1}, \cdots, \hat{u}_{T-1}) \\ &\leq \sup\nolimits_{\in U_k} \left[f_k^0(x,u) + V_{k+1}(f_k(x,u))\right], \end{aligned}$$

de donde se deduce, desde que ϵ es abitrario, que

$$V_k(x) \leq \sup_{u \in U_k} \left[f_k^0(x,u) + V_{k+1}(f_k(x,u))\right].$$

Por lo que, juntando ambas desigualdades, hemos probado la igualdad. ■

Corolario 2.2.1 (Caracterización de las soluciones) *Una sucesión de decisiones* $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ *con su proceso asociado* $(\hat{x}_0, \cdots, \hat{x}_T)$ *será una solución óptima del problema* $\mathcal{P}_T(\hat{x}_0)$ *si, y sólo si, satisfacen las ecuaciones de Bellman. Esto es, para todo* $t = 0, \cdots, T-1$,

$$\begin{aligned} V_t(\hat{x}_t) &= \max_{u \in U_t}\{f_t^0(\hat{x}_t, u) + V_{t+1}(f_t(\hat{x}_t, u))\} \\ &= f_t^0(\hat{x}_t, \hat{u}_t) + V_{t+1}(f_t(\hat{x}_t, \hat{u}_t)). \end{aligned}$$

Prueba. $(\Rightarrow)$ Tomemos $t \in \{0, \cdots, T-2\}$ fijo y arbitrario. Del Principio de optimalidad de Bellman, tenemos que $(\hat{u}_{t+1}, \hat{u}_{t+2}, \cdots, \hat{u}_{T-1})$ es solución óptima del problema $\mathcal{P}_T^{t+1}(\hat{x}_{t+1})$. Por ello se cumple por su definición

$$V_{t+1}(f_t(\hat{x}_t, \hat{u}_t)) = V_{t+1}(\hat{x}_{t+1}) = B_T^{t+1}(\hat{x}_{t+1}; \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}). \qquad (2.1)$$

Similarmente $(\hat{u}_t, \hat{u}_{t+1}, \cdots, \hat{u}_{T-1})$ es solución óptima del problema $\mathcal{P}_T^t(\hat{x}_t)$. Por ello se cumple de su definición

$$V_t(\hat{x}_t) = B_T^t(\hat{x}_t; \hat{u}_t, \cdots, \hat{u}_T) = f_t^0(\hat{x}_t, \hat{u}_t) + B_T^{t+1}(\hat{x}_{t+1}; \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}).$$

Usando (2.1) obtenemos para $t = 0, \cdots, T-2$,

$$V_t(\hat{x}_t) = f_t^0(\hat{x}_t, \hat{u}_t) + V_{t+1}(f_t(\hat{x}_t, \hat{u}_t)). \qquad (2.2)$$

Cuando $t = T-1$, entonces $\hat{u}_{T-1}$ es solución del problema $\mathcal{P}_T^{T-1}(\hat{x}_{T-1})$ y por lo tanto

$$V_{T-1}(\hat{x}_{T-1}) = f_{T-1}^0(\hat{x}_{T-1}, \hat{u}_{T-1}) + f_T^0(\hat{x}_T).$$

Haciendo uso la convención $V_T \equiv f_T^0$, tenemos que (2.2) también se verifica para $t = T-1$.

$(\Leftarrow)$ Recíprocamente, si un camino cumpliese para todo $t = 0, \cdots, T-1$,

$$V_t(\hat{x}_t) = f_t^0(\hat{x}_t, \hat{u}_t) + V_{t+1}(f_t(\hat{x}_t, \hat{u}_t)),$$

entonces es fácil probar inductivamente que para cada $t = 0, \cdots, T-1$,

$$V_0(\hat{x}_0) = \sum_{s=0}^{t} f_s^0(\hat{x}_s, \hat{u}_s) + V_{t+1}(\hat{x}_{t+1}).$$

Así, para $t = T-1$, y del hecho $V_T(\hat{x}_T) = f_T^0(\hat{x}_T)$, obtenemos

$$V_0(\hat{x}_0) = \sum_{t=0}^{T-1} f_t^0(\bar{x}_t, \bar{u}_t) + f_T^0(x_T).$$

Debido a que $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ es admisible, se deduce que este es una solución óptima de $P_T(\hat{x}_0)$. ■

En base al resultado anterior obtenemos un método general que resuelve "todos" los problemas de control, siempre que puedan calcularse cada paso que se necesite. Describamos a continuación este método, como un algoritmo.

2.3 Algoritmo de Bellman

Algoritmo

1. **Inicialización :** Asumir $V_T = f_T^0$, hacer $k = 1$.

2. **Etapa k :** Para cada $x \in X_{T-k}$ calcular

$$V_{T-k}(x) = \max_{u \in U_{T-k}} \{f_{T-k}^0(x,u) + V_{T-k+1}(f_{T-k}(x,u))\},$$

   y elegir

$$d_{T-k} : X_{T-k} \to U_{T-k}$$

   definida por

$$d_{T-k}(x) = u, \text{ tal que } V_{T-k}(x) = f_{T-k}^0(x,u) + V_{T-k+1}(f_{T-k}(x,u)).$$

   Hacer $k = k + 1$. Retornar al ítem 2.

Comentarios. En general la función d_j del algoritmo puede no existir, en caso existiese, puede no ser única.

Un caso en el que podemos aplicar este algoritmo es cuando para $t = 0, \cdots, T-1$, los conjuntos X_{t+1}, U_t son finitos. En este caso es posible construir punto por punto las funciones V_t y d_t.

Un caso más general en el que puede aplicarse este algoritmo lo damos en el siguiente lema.

Lema 2.3.1 (Aplicación del Algoritmo) *Asumamos las mismas condiciones del Teorema 1.2.2, entonces para $k = 0, \cdots, T-1$, las funciones V_k son continuas y siempre es posible construir las funciones d_k para cada $x_k \in X_k$. Por lo tanto el algoritmo es ejecutable.*

Prueba. Para probar el lema probemos primero una afirmación

Afirmación *Asumamos que $X \subseteq \mathbb{R}^n$ y $U \subseteq \mathbb{R}^m$ son no vacíos con U compacto y $g : X \times U \to \mathbb{R}$ continua. Entonces la función $f : X \to \mathbb{R}$ definida por*

$$f(x) = \max_{u \in U} g(x,u),$$

es continua sobre X.

En efecto, asociado a una función arbitraria $h : \mathbb{R}^m \supset Z \to \mathbb{R}$, definamos los conjuntos

$$\widetilde{\text{epi}}\,(h) := \{(z,\lambda) \in Z \times \mathbb{R} : h(z) < \lambda\}$$

y

$$\widehat{\text{epi}}\,(h) := \{(z,\lambda) \in Z \times \mathbb{R} : h(z) > \lambda\}.$$

Se sabe que h es continua sobre Z si, y sólo si, $\widetilde{\text{epi}}\,(h)$ y $\widehat{\text{epi}}\,(h)$ son abiertos en $Z \times \mathbb{R}$.

Por definición,

$$\widetilde{\text{epi}}\,(f) = \text{proj}_{\,X\times\mathbb{R}}(\widetilde{\text{epi}}\,(g)) \quad \text{y} \quad \widehat{\text{epi}}\,(f) = \text{proj}_{\,X\times\mathbb{R}}(\widehat{\text{epi}}\,(g)).$$

Debido a que g es continua, $\widetilde{\text{epi}}\,(g)$ y $\widehat{\text{epi}}\,(g)$ son abiertos y por lo tanto sus respectivas proyecciones $\widetilde{\text{epi}}\,(f)$ y $\widehat{\text{epi}}\,(f)$, también son abiertos. Se sigue que f es continua sobre X. Esta función es finita sobre X debido a que U es compacto no vacío.

Usando la afirmación anterior, probemos por recursión que las funciones V_k son continuas. Tenemos, por convención, que $V_T \equiv f_T^0$ por lo cual es continua. De otro lado definamos la función $g_{T-1} : X_{T-1} \times U_{T-1} \to \mathbb{R}$ por

$$g_{T-1}(x,u) = f_{T-1}^0(x,u) + V_T(f_{T-1}(x,u)).$$

Por hipótesis, g_{T-1} es continua y, debido a que U_{T-1} es compacto, la función

$$V_{T-1}(x) = \max_{u\in U_{T-1}} (f_{T-1}^0(x,u) + V_T(f_{T-1}(x,u))) = \max_{u\in U_{T-1}} g_{T-1}(x,u).$$

también es continua sobre X_{T-1}. Procediendo de modo similar, podemos demostrar recursivamente que las funciones V_k son continuas para $k = 0, \cdots, T-1$.

Debido a que U_k es compacto, por el teorema de Weierstrass, para todo $x_k \in X_k$ existe $u_k \in U_k$ tal que

$$V_k(x_k) = f_k^0(x_k,u_k) + V_{k+1}(f_k(x_k,u_k)).$$

Por ello podemos tomar $d_k(x_k) = u_k$. Como x_k es arbitrario, deducimos que podemos en principio construir la función d_k en todo X_k. Por ello el algoritmo es realizable. ■

Haciendo uso de este algoritmo, podemos encontrar (compare con el Teorema 1.2.2) una solución óptima del problema $\mathcal{P}_T(\hat{x}_0)$ desde que las funciones d_k $(k = 0, \cdots, T-1)$ pueden ser determinadas. En efecto, el proceso $\hat{x}_0, \cdots, \hat{x}_T$ con sus decisiones $\hat{u}_0, \cdots, \bar{u}_{T-1}$, con $\hat{u}_k = d_k(\hat{x}_k)$ y $\hat{x}_{k+1} = f_{k+1}(\hat{x}_k, \hat{u}_k)$, cumplen las ecuaciones de Bellman.

Con esta metodología no sólo hemos hallado la solución del problema $\mathcal{P}_T(\hat{x}_0)$ sino también para cada uno de los problemas de la familia $(\mathcal{P}_T(x))_{x\in X_0}$. Más aún, tenemos las soluciones para todos los problemas $(\mathcal{P}_T^k(x))_{x\in X_k}$.

Corolario 2.3.1 *Para $k = 0, \cdots, T-1$, dado $\hat{x}_k \in X_k$, la sucesion de decisiones $\hat{u}_k, \cdots, \hat{u}_{T-1}$, definida recursivamente para $t = k, \cdots, T-2$,*

$$\hat{u}_{t+1} := d_{t+1}(f_t(\hat{x}_t, d_t(\hat{x}_t))),$$

donde $\hat{u}_k := d_k(\hat{x}_k)$, es una solución del problema $\mathcal{P}_T^k(\hat{x}_k)$.

Esta propiedad de la metodología desarrollada es al menos teóricamente muy satisfactoria, pues en esencia hemos resuelto completamente el problema de control discreto. Sin embargo este método es "ineficiente" pues exige implícitamnte ir resolviendo **todos** los problemas de menor orden antes de arribar a una solución para nuestro problema original $\mathcal{P}_T(\hat{x}_0)$, lo cual para fines prácticos es en cierta medida inaplicable.

Tenemos pues la necesidad de proponer metodologías diferentes que permitan obtener el resultado, tal vez más rápidamente, o que combinadas con la programación dinámica nos acerquen a la solución de $\mathcal{P}_T(\hat{x}_0)$ más eficientemente. Una forma de conseguir esto sería encontrar una condición necesaria de optimalidad, calcular los puntos que la satisfacen, y luego testear con el sistema de Bellman para ver si son óptimas. Por esta razón estudiaremos en los dos capítulos siguientes el Principio del Máximo de Pontryagin en su versión discreta.

Capítulo 3

Principio Débil de Pontryagin Discreto

En el presente capítulo construímos la versión discreta del Principio del Máximo de Pontryagin, al que denominamos Principio Fuerte de Pontryagin Discreto (PFP). A su vez definimos lo que se entenderá por Principios Débiles de Pontryagin Discretos (PdP).

El aporte central de este capítulo es haber establecido sistemáticamente los diferentes contextos, crecientes en generalidad, en los que cada Principio Débil es una condición necesaria de optimalidad. Confirmando con esto que el Principio del Máximo de Pontryagin, también se cumple para el problema discreto, pero "débilmente."

En la Teoría de Control Continua, el Principio del Máximo es una condición necesaria para la optimalidad. En el caso convexo éste también es una condición suficiente.

Debido a la gran utilidad que tiene este principio en la solución de varios problemas concretos, es importante tratar de encontrarle una versión discreta.

Con esta intención un primer paso para encontrar esta versión será discretizar este principio partiendo del caso continuo para obtener heurísticamente una versión tentativa que nos servirá para nuestro fin. Recordemos entonces la formulación continua.

Sean $f : \Omega \times U \to \mathbb{R}^n$, $f^0 : \Omega \times U \to \mathbb{R}$, $f_T^0 : \Omega \to \mathbb{R}$, donde $T > 0$, $\Omega \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$ y $t \in [0, T]$. Para $\hat{x}_0 \in \Omega$ consideremos el siguiente problema de control continuo en horizonte finito T.

$$\begin{array}{ll} \max & B(x(\cdot), u(\cdot)) := \int_0^T f^0(x(t), u(t))dt + f_T^0(x_T), \\ \text{s.a.} & (x(\cdot), u(\cdot)) \in \mathrm{Adm}\,(\hat{x}_0), \end{array} \qquad (\mathcal{PC}_T(\hat{x}_0))$$

donde,

$$\operatorname{Adm}(\hat{x}_0) := \left\{ (x(\cdot), u(\cdot)) \,\middle|\, \begin{array}{l} u : [0,T] \to U,\ x : [0,T] \to \Omega,\ \forall\, t \in [0,T], \\ \dot{x}(t) = f(x(t), u(t)),\ x(0) = \hat{x}_0,\ x(T) \in M_f \end{array} \right\},$$

con $M_f \subseteq \mathbb{R}^n$.

Para simplificar el enunciado del Principio del Máximo de Pontryaguin, asumamos que $M_f = \mathbb{R}^n$ e introduzcamos los siguientes elementos.

Hamiltoniano Asociado al problema $\mathcal{PC}_T(\hat{x}_0)$, se define el Hamiltoniano como la función $H : (\mathbb{R}^n)^* \times \Omega \times U \to \mathbb{R}$,

$$H(\varphi, x, u) := f^0(x, u) + \langle \varphi\ ,\ f(x, u) \rangle. \tag{H}$$

Donde $(\mathbb{R}^n)^*$ denota el espacio dual de $\mathbb{R}^n$. El producto $\langle \cdot, \cdot \rangle$ denota el producto dual sobre $(\mathbb{R}^n)^* \times \mathbb{R}^n$.

La definición (H) también puede ser escrita como

$$H(\varphi, x, u) = \sum_{i=0}^{n} \varphi_i f^i(x, u),$$

donde $\varphi_0 = 1, \varphi = (\varphi_1, \cdots, \varphi_n), f(x, u) = (f^1(x, u), \cdots, f^n(x, u))$. Aquí las componentes de la funcional φ se interpretan como números reales.

Variable Adjunta Asociado al control $u(\cdot)$ y su correspondiente camino $x(\cdot)$, la variable adjunta es la función $\varphi = (\varphi_1, \cdots, \varphi_n) : [0,T] \to (\mathbb{R}^n)^*$ satisfaciendo para $j = 1, \cdots, n$,

$$\frac{d\varphi_j}{dt} = -\frac{\partial H}{\partial x_j}(\varphi, x, u) = -\sum_{t=0}^{n} \varphi_i \frac{\partial f^i}{\partial x_j}(x, u), \tag{VA}$$

con la condición final,

$$\varphi_j(T) = \frac{\partial f_T^0}{\partial x_j}(x_T) \qquad \text{y} \qquad \varphi_0 = 1. \tag{CF}$$

Teorema 3.0.1 (Principio de Máximo de Pontryagin Continuo) *Sea la curva $(\bar{x}(\cdot), \bar{u}(\cdot))$ una solución del problema $\mathcal{PC}_T(\hat{x}_0)$, con $M_f = \mathbb{R}^n$, y sea $\bar{\varphi}(\cdot)$ su variable adjunta. Entonces para todo $t \in [0,T]$,*

$$H(\bar{\varphi}(t), \bar{x}(t), \bar{u}(t)) = \max_{u \in U} H(\bar{\varphi}(t), \bar{x}(t), u).$$

Prueba. Ver [7, p.68]. ∎

3.1 Principio Tentativo de Pontryagin Discreto

Usaremos el Principio del Máximo de Pontryagin enunciado en el Teorema 3.0.1 para encontrar una versión discreta, con la cual formularemos una condición necesaria de optimalidad. En el Capítulo 4 estudiaremos bajo qué condiciones ésta será válida y daremos contraejemplos que muestran que en general no es cierta.

Para formular este principio tentativo realizemos los siguientes 8 pasos.

(1) **Partición del Intervalo** $[0, T]$**.** Dividamos el intervalo $[0, T]$ en N tiempos iguales, obtenemos la partición $\{0, h, 2h, \cdots, Nh = T\}$, donde $h := \frac{T}{N}$. Interpretaremos estos elementos como los tiempos $t \in \{0, 1, \cdots, N\}$.

(2) **Discretización de la Dinámica.** Asociado al control $u : [0, T] \to U$, y a su camino correspondiente $x : [0, T] \to \Omega$, denotemos para cada $t = 0, 1, \cdots, N$,

$$u_t = u(th) \in U \qquad \text{y} \qquad x_t = x(th) \in \Omega.$$

Con estos elementos aproximamos la dinámica por

$$\frac{x_{t+1} - x_t}{h} \approx \frac{dx}{dt}(th) \approx f(x_t, u_t),$$

obteniéndose la discretización para $t = 0, 1, \cdots, N - 1$,

$$x_{t+1} = x_t + hf(x_t, u_t),$$

con las restricciones $x_t \in \Omega,\ u_t \in U \subseteq \mathbb{R}^m$ y $x_0 := x(0) = \hat{x}_0$ (punto inicial).

Definiendo para cada $t = 0, \cdots, N - 1$, las funciones $f_t : \Omega \times U \to \mathbb{R}^n$, como $f_t(x_t, u_t) := x_t + h.f(x_t, u_t)$. Tenemos la discretización final de la dinámica para $t = 0, 1, \cdots, N - 1$,

$$x_{t+1} = f_t(x_t, u_t).$$

(3) **Discretización de la Función Objetivo.** Teniendo en cuenta las discretizaciones anteriores, la función objetivo queda expresada en forma discreta como:

$$B_T(x_0, x_1, \cdots, x_N, u_0, \cdots, u_{N-1}) := \sum_{t=0}^{N-1} hf^0(x_t, u_t) + f_T^0(x_N).$$

(4) **Discretización de las Variables Adjuntas (VA).** Asociados a la variable adjunta $\varphi(\cdot)$ del camino $(x(\cdot), u(\cdot))$, definamos

$$\psi(t) := \begin{cases} \varphi(th), & \text{si } t = 0, 1, \cdots, N-2, \\ \varphi(T), & \text{si } t = N-1. \end{cases}$$

De esto tenemos

$$H(\psi(t), x_t, u_t) = f^0(x_t, u_t) + \langle \psi(t); f(x_t, u_t) \rangle = \sum_{i=0}^{n} \psi_i(t) f^i(x_t, u_t),$$

donde $\psi_0(t) := 1$.

Con respecto a las ecuaciones **(VA)** podemos aproximarlas por

$$\begin{aligned} \tfrac{\psi_j(t)-\psi_j(t+1)}{-h} &\approx \tfrac{d\varphi_j}{dt}((t+1)h) \approx -\tfrac{\partial H}{\partial x^j}(\psi(t+1), x_{t+1}, u_{t+1}) \\ &= -\textstyle\sum_{i=0}^{n} \psi_i(t+1) \tfrac{\partial f^i}{\partial x^j}(x_{t+1}, u_{t+1}), \end{aligned}$$

de donde para $j = 1, \cdots, n$, se obtiene

$$\psi_j(t) \approx \psi_j(t+1) + h \sum_{i=0}^{n} \psi_i(t+1) \frac{\partial f^i}{\partial x^j}(x_{t+1}, u_{t+1}), \qquad (EAdis)$$

para cada $t = 0, \cdots, N-2$.

Con condición final

$$\psi_j(N-1) = \frac{\partial f_T^0}{\partial x^j}(x_T).$$

(5) **Definición de las Funciones.** Para que nuestra notación sea más acorde con la usada para el problema $\mathcal{P}_T(\hat{x}_0)$, definamos las funciones de beneficio y evolución. Para $t = 0, \cdots, N-1$,

$$f_t^0(x_t, u_t) := h.f^0(x_t, u_t) \qquad \text{y} \qquad f_t(x_t, u_t) := x_t + h.f^0(x_t, u_t).$$

Definamos además los Hamiltonianos asociados como

$$\begin{aligned} H_t(x_t, u_t, \psi(t)) &= f_t^0(x_t, u_t) + \langle \psi(t); f_t(x_t, u_t) \rangle \\ &= h.f^0(x_t, u_t) + \langle \psi(t); x_t + h.f(x_t, u_t) \rangle. \end{aligned}$$

(6) **Definición del Problema Discreto.** Con las notaciones anteriores tenemos la siguiente discretización del problema $(\mathcal{PC}_T(\hat{x}_0))$,

$$\begin{array}{ll} \max & B_T(x_0, \cdots, x_N, u_0, \cdots, u_{N-1}) = \sum_{t=0}^{T-1} f_t^0(x_t, u_t) + f_T^0(x_T) \\ \text{s.a} & \forall\, t = 1, \cdots, T, \; x_t \in X_t, \; u_{t-1} \in U_{t-1}, \\ & \forall\, t = 1, \cdots, T, \; x_{t+1} = f_t(x_t, u_t), \\ & x_0 = \hat{x}_0 \in X_0, \end{array} \qquad (\mathcal{P}_T^c(\hat{x}_0))$$

donde para $t = 0, \cdots, T-1$,

- $X_t := \Omega$, $U_t := U$ y $X_T := \Omega$.
- $f_t^0 : \Omega \times U \to \mathbb{R}$, $f_t : \Omega \times U \to \mathbb{R}^n$ y $f_T^0 : \Omega \to \mathbb{R}$ son funciones definidas en el paso anterior.

Notemos que se ha denotado a este problema con superíndice c para dejar claro que proviene del caso continuo y no confundirlo con la representación general $\mathcal{P}_T(\hat{x}_0)$. Sin embargo notamos que es un caso particular de éste.

(7) **Definición del Sistema Adjunto Discreto (EAD).** Teniendo en cuenta la redefinición en la etapa (5) y el sistema (EAdis), el sistema adjunto correspondiente al problema $\mathcal{P}_T^c(\hat{x}_0)$ viene dado por:

Para $j = 1, \cdots, n$,

$$\psi_j(t) = \sum_{i=0}^{n} \psi_i(t+1) \frac{\partial f_{t+1}^i}{\partial x^j}(x_{t+1}, u_{t+1}),$$

$$\psi_0(t) = 1, \qquad\qquad t = 0, \cdots, T-1,$$

con la condición final

$$\psi_j(N-1) = \frac{\partial f_T^0}{\partial x^j}(x_T).$$

Ahora, teniendo en cuenta la definición de los hamiltonianos asociados en el paso (5), se deduce

$$\sum_{i=0}^{n} \psi_i(t+1) \frac{\partial f_{t+1}^i}{\partial x^j}(x_{t+1}, u_{t+1}) = \frac{\partial H_{t+1}}{\partial x^j}(\psi(t+1), x_{t+1}, u_{t+1}).$$

De esto, el anterior sistema adjunto discretizado queda expresado como.

Para $t = 1, \cdots, N-1$,

$$\psi(t-1) = \nabla_x H_t(\psi(t), x_t, u_t), \qquad (EAD)$$

con condición final

$$\psi(N-1) = \nabla_x(f_T^0(x_T)).$$

(8) **Variables Adjuntas** Dada una sucesión de decisiones $(u_0, \cdots, u_{N-1})$ y su proceso correspondiente $(x_0, \cdots, x_N)$ existe una colección de vectores $\{\psi(0), \cdots, \psi(N-1)\} \subset \mathbb{R}^n$ que solucionan el sistema (EAD). A cada uno de estos vectores se les denominará variable adjunta.

Debido a la condición final, tal colección es única.

Ahora estamos en condición de poder formular lo que llamaremos Principio tentativo de Pontryaguin discreto. Lo llamamos tentativo porque lo hemos obtenido a partir de la versión continua y aún no sabemos si es válido.

Proposición 3.1.1 (Principio Tentativo de Pontryagin Discreto) *Asumamos que la sucesión de decisiones* $(\bar{u}_0, \cdots, \bar{u}_{N-1})$ *con su proceso asociado* $(\bar{x}_0, \cdots, \bar{x}_N)$ *es una solución óptima del problema* $\mathcal{P}_T^c(\hat{x}_0)$.

Entonces existen vectores $\bar{\psi}(0), \cdots, \bar{\psi}(N-1) \in \mathbb{R}^n$ *tales que:*

1. *Para* $t = 1, \cdots, N-1$,

$$\bar{\psi}(t-1) = \nabla_x(H(\bar{\psi}(t), \bar{x}_t, \bar{u}_t), \qquad (EAD)$$

con condición final $\bar{\psi}(N-1) = \nabla_x f_T^0(\bar{x}_N)$.

2. *Para cada* $t = 0, \cdots, N-1$, $\bar{u}_t$ *es una solución del problema*

$$\max_{u \in U_t} H(\bar{\psi}(t), \bar{x}_t, u). \qquad (PMD)_t$$

Prueba. En estudio. ■

El principio tentativo enunciado anteriormente nos lleva a extenderlo para el problema discreto $\mathcal{P}_T(\hat{x}_0)$. Este principio será llamado **Principio Fuerte de Pontryagin (PFP)** y como veremos en el Capítulo 4 éste será válido para ciertos casos especiales, es decir, en general la afirmación no es cierta.

Asociado al problema $\mathcal{P}_T(\hat{x}_0)$ para cada $t = 0, \cdots, T-1$, los hamiltonianos para este problema se definirán como las funciones $H_t : X_t \times U_t \times \mathbb{R}^n \to \mathbb{R}$ tales que $H_t(x, u, \phi) = f_t^0(x, u) + \langle \phi, f_t(x, u) \rangle$.

Proposición 3.1.2 (Principio Fuerte de Pontryagin (PFP)) *Para el problema* $\mathcal{P}_T(\hat{x}_0)$ *sea la sucesión de decisiones* $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, *con su respectivo proceso asociado* $(\hat{x}_0, \cdots, \hat{x}_T)$ *una solución óptima.*

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$ *tales que*

1. *Para* $t = 1, \cdots, T-1$,

$$\hat{\psi}_{t-1} = \nabla_x(H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \qquad (EA)$$

con condición final $\hat{\psi}_{T-1} = \nabla_x f_T^0(\hat{x}_T)$.

2. *Para cada* $t = 0, \cdots, T-1$, $\mathbf{\hat{u}(t)}$ *es una solución óptima del problema*

$$\max_{u \in U_t} H_t(\hat{x}_t, u, \hat{\psi}_t). \qquad (PM)_t$$

Llamamos “Fuerte” a este principio porque exige que las decisiones óptimas $(\hat{u}_t)$ sean, a su vez, soluciones óptimas de los subproblemas $(PM)_t$. Debilitando esta condición, tendremos teoremas en los que sólo se exige que las decisiones $(\hat{u}_t)$ sean puntos críticos para estos problemas $(PM)_t$ (en el sentido de que satizfacen alguna condición necesaria de óptimo). Los teoremas de este último tipo serán denominados “Principios Débiles”.

El Principio Fuerte de Pontryagin ¿Será válido? Veremos que para ciertos casos especiales sí lo es. En general la afirmación no es cierta. Existen contra-ejemplos. No obstante en contextos bastante generales se cumplen teoremas a los que llamamos “Principios Débiles”.

Principios Débiles de Pontryagin Llamaremos “Principios Débiles de Pontryagin” a aquellos teoremas que son condiciones necesarias de optimalidad para procesos discretos y que tienen como esquema general la siguiente formulación:

Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{N-1})$ *óptimas para* $\mathcal{P}_T(\hat{x}_0)$, *con su proceso asociado* $(\hat{x}_0, \cdots, \hat{x}_N)$.

Entonces para $t = 0, \cdots, N-1$, *existen vectores* $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, *a los que se les llama “adjuntos” que satisfacen:*

1. *Un sistema de ecuaciones (sistema adjunto), con su condición final respectiva.*

2. *Para cada* $t = 0, \cdots, N-1$,

$\hat{u}_t$ **es un punto crítico** *(es decir, satisface alguna condición necesaria de optimalidad: Lagrange, KKT, Clarke, etc.) para el problema*

$$\max_{u \in U_t} H_t(\hat{x}_t, u, \hat{\psi}_t),$$

donde H_t *representa el Hamiltoniano en el tiempo “t” del problema, cualesquiera sea sus forma o generalidad.*

Claramente los teoremas que satisfagan este esquema de representación “debilitan” la Proposición 3.1.2 que hemos llamado **Principio Fuerte de Pontryagin (PFP)**.

Estudiemos primero estos principios débiles, por ser verdaderos en contextos más generales que el principio fuerte. En esta dirección recordemos el teorema de existencia de solución, demostrado en el Capítulo 1. Hemos visto como podíamos transformar nuestro problema dinámico en uno estático. La idea principal para demostrar todos estos principios se basa en usar esta transformación y luego al problema equivalente aplicarle alguna condición necesaria de óptimo, de la teoría de optimización estática. Para hacer énfasis en este punto tenemos el siguiente lema.

Lema 3.1.1 (Control Dinámico a Optimización Estática) *El problema de control $\mathcal{P}_T(\hat{x}_0)$ es equivalente a un problema de optimización estática.*

Prueba. Para $t = 0, \cdots, T$ y $s = 0, \cdots, T-1$, denotemos los vectores $x_t = (x_t^1, \cdots, x_t^n) \in X_t \subseteq \mathbb{R}^n$ y $u_s = (u_s^1, \cdots, u_s^m) \in U_t \subseteq \mathbb{R}^m$, y la función $f_t = (f_t^1, \cdots, f_t^n) : X_t \times U_t \to \mathbb{R}^n$. Denotemos $Z := \prod_{t=1}^{T} X_t \times \prod_{s=0}^{T-1} U_s \subseteq \mathbb{R}^{(n+m)T}$ y a sus elementos como $z = (x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$.

Definamos $F^0, F_t^i : Z \to \mathbb{R}$ por

$$F^0(z) = \sum_{k=0}^{T-1} f_k^0(x_k, u_k) + f_T^0(x_T) \quad \text{ y } \quad F_t^i(z) = x_{t+1}^i - f_t^i(x_t, u_t),$$

donde $x_0 = \hat{x}_0$.

Asociados a estas funciones definimos el siguiente conjunto

$$\Omega := \left\{ z \in Z \;\middle|\; F_t^i(z) = 0, \text{ para } i = 1, \cdots, n, \text{ y } \quad t = 0, \cdots, T-1 \right\}.$$

Así, con estos elementos, el problema $\mathcal{P}_T(\hat{x}_0)$ es equivalente al problema estático

$$\max_{z \in \Omega} F^0(z).$$

Explícitamente, el vector $(\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ es solución de $\mathcal{P}_T(\hat{x}_0)$ si, y sólo si es solución de $\max_{z \in \Omega} F^0(z)$. ■

3.2 Principio Débil: Caso Interior

El objetivo de esta sección es demostrar una condición de optimalidad para el problema de control discreto $\mathcal{P}_T(\hat{x}_0)$ usando su equivalencia estática mostrada en el Lema 3.1.1 al que aplicaremos el Teorema de Multiplicadores de Lagrange. Para poder aplicar correctamente este teorema es necesario exigir ciertas condiciones sobre las funciones. Estas hipótesis son:

1. Para $t = 0, \cdots, T-1$,

$$f_t^0 \text{ es diferenciable en } X_t \times U_t \quad \text{ y } \quad f_T^0 \in C^1(X_T). \tag{$H1$}$$

2. Para $t = 0, \cdots, T-1$,

$$f_t \in C^1(X_t \times U_t). \tag{$H2$}$$

Bajo estas dos hipótesis probemos el siguiente lema que será esencial para poder usar el Teorema de Multiplicadores de Lagrange.

En lo que sigue siempre que hablemos de las funciones F_k^i, nos referiremos a las funciones que hemos definido en el Lema 3.1.1. Observemos que estas funciones están asociadas al problema $\mathcal{P}_T(\hat{x}_0)$ y no dependen de las hipótesis adicionales que se hagan sobre éste.

Lema 3.2.1 *Asumamos la hipótesis $(H2)$, entonces las funciones F_k^i son diferenciables en Z y para cada $z \in Z$ el conjunto*

$$\left\{\nabla_z F_k^i(z)\right\}_{\substack{i=1,\cdots,n,\\ k=0,\cdots,T-1.}} \subseteq \mathbb{R}^{(n+m)T}$$

es linealmente independiente

Prueba. Debido a que $F_k^i(z) = x_{k+1}^i - f_k^i(x_k, u_k)$, la hipótesis $(H2)$ implica que estas funciones son de clase C^1.

1. Cálculo de los Gradientes. En general se tiene

$$\nabla_z F_k^i(z) = \left(\nabla_{x_1} F_k^i(z), \cdots, \nabla_{x_T} F_k^i(z), \nabla_{u_0} F_k^i(z), \cdots, \nabla_{u_{T-1}} F_k^i(z)\right),$$

donde $\nabla_{x_t} F_k^i(z) \in \mathbb{R}^n$ y $\nabla_{u_t} F_k^i(z) \in \mathbb{R}^m$. Consideremos $k \in \{0, \cdots, T-1\}$ fijo arbitrario. Debido a que $F_k^i(z) = x_{k+1}^i - f_k^i(x_k, u_k)$ depende solo de x_{k+1}, x_k y u_k (cuando $k = 0$, solo depende de x_1 y u_1, pues $x_0 = \hat{x}_0$) deducimos:

- Si $s \notin \{k, k+1\}$, $\nabla_{x_s} F_k^i(z) = 0$.
- Si $r \neq k$, $\nabla_{u_r} F_k^i(z) = 0$.
- Si $s = k+1$,

  $$\frac{\partial F_k^i}{\partial x_{k+1}^j}(z) = \delta_j^i \Rightarrow \nabla_{x_{k+1}} F_k^i(z) = e_i \text{ (elemento de la base canónica de } \mathbb{R}^n).$$

- Si $s = k$, con $k > 0$,

  $$\frac{\partial F_k^i}{\partial x_k^j}(z) = -\frac{\partial f_k^i}{\partial x_k^j}(x_k, u_k) \Rightarrow \nabla_{x_k} F_k^i(z) = -\nabla_{x_k} f_k^i(x_k, u_k).$$

- Similarmente, si $r = k$,

  $$\nabla_{u_k} F_k^i(z) = -\nabla_{u_k} f_k^i(x_k, u_k).$$

Resumiendo tenemos,

$$\nabla_z F_0^i(z) = (\underline{e_i}_{\,1}, \cdots, 0, \cdots, \underline{0}_{\,T}; \underline{-\nabla_{u_0} f_0^i(x_0, u_0)}_{\,0}, 0, \cdots, \underline{0}_{\,T-1})$$

donde cada elemento es un vector de $\mathbb{R}^n$ y, para $k = 1, \cdots, T-1$,

$$\begin{aligned}\nabla_z F_k^i(z) \quad = \quad & \Big(\underline{0}_{\,1}, \cdots, 0, \underline{-\nabla_{x_k} f_k^i(x_k, u_k)}_{\,k}, \underline{e_i}_{\,k+1}, 0, \cdots, \underline{0}_{\,T}; \\ & ; \underline{0}_{\,0}, \cdots, 0, \underline{-\nabla_{u_k} f_k^i(x_k, u_k)}_{\,k}, 0, \cdots, \underline{0}_{\,T-1}\Big).\end{aligned}$$

2. Independencia Lineal. Dado $z \in Z$ fijo arbitrario y, para $i = 1, \cdots, n$ y $k = 1, \cdots, T-1$, consideremos $\lambda_k^i \in \mathbb{R}$ tal que

$$\sum_{\substack{i=1,\cdots,n \\ k=1,\cdots,T-1}} \lambda_k^i \nabla_z F_k^i(z) = 0.$$

Entonces

$$0 = \sum_{\substack{i=1,\cdots,n \\ k=1,\cdots,T-1}} \lambda_k^i (\nabla_{x_1} F_k^i(z), \cdots, \nabla_{x_T} F_k^i(z); \nabla_{u_0} F_k^i(z), \cdots, \nabla_{u_{T-1}} F_k^i(z)).$$

De lo cual, para $s = 1, \cdots, T$ y $t = 0, \cdots, T-1$,

$$\sum_{i,k} \lambda_k^i \nabla_{x_s} F_k^i(z) = 0 \quad \text{y} \quad \sum_{i,k} \lambda_k^i \nabla_{u_t} F_k^i(z) = 0.$$

Considerando los cálculos de los gradientes (item 1), para $s = T$, tenemos

$$0 = \sum_{\substack{i=1,\cdots,n \\ k=0,\cdots,T-1}} \lambda_k^i \nabla_{x_T} F_k^i(z) = \sum_{i=1,\cdots,n} \lambda_{T-1}^i \nabla_{x_T} F_{T-1}^i(z) = \sum_{i=1,\cdots,n} \lambda_{T-1}^i e_i,$$

de donde

$$(\lambda_{T-1}^1, \cdots, \lambda_{T-1}^n) = 0.$$

A continuación tomemos $s = T-1$. Usando los cálculos de los gradientes y el resultado anterior para $s = T$, tenemos

$$\begin{aligned} 0 &= \sum_{\substack{i=1,\cdots,n \\ k=0,\cdots,T-1}} \lambda_k^i \nabla_{x_{T-1}} F_k^i(z) = \sum_{\substack{i=1,\cdots,n \\ k=T-2,T-1}} \lambda_k^i \nabla_{x_{T-1}} F_k^i(z) \\ &= \textstyle\sum_{i=1,\cdots,n} \lambda_{T-2}^i \nabla_{x_{T-1}} F_{T-2}^i(z) = \sum_{i=1,\cdots,n} \lambda_{T-2}^i e_i. \end{aligned}$$

De este se deduce

$$(\lambda_{T-2}^1, \cdots, \lambda_{T-2}^n) = 0.$$

Por inducción, supongamos que hasta para $r = s, \cdots, T-1$,

$$(\lambda_r^1, \cdots, \lambda_r^n) = 0,$$

estudiemos para $r = s-1$. Sabemos que si $k \notin \{s-1, s\}$, $\nabla_{x_s} F_k^i(z) = 0$, por lo que

$$\begin{aligned} 0 &= \sum_{\substack{i=1,\cdots,n \\ k=0,\cdots,T-1}} \lambda_k^i \nabla_{x_s} F_k^i(z) = \sum_{\substack{i=1,\cdots,n \\ k=s-1,s}} \lambda_k^i \nabla_{x_s} F_k^i(z) \\ &= \textstyle\sum_{i=1,\cdots,n} \nabla_{x_s} F_{s-1}^i(z) = \sum_{i=1,\cdots,n} \lambda_{s-1}^i e_i. \end{aligned}$$

De donde

$$(\lambda_{s-1}^1, \cdots, \lambda_{s-1}^n) = 0.$$

Procediendo de esta forma obtenemos que necesariamente para $r = 0, \cdots, T-1$,

$$(\lambda_r^1, \cdots, \lambda_r^n) = 0.$$

Como esta combinación es arbitraria deducimos que el conjunto de gradientes de las funciones F_k^i en $z \in Z$ es lineal independiente. Debido a que z también es arbitrario deducímos el teorema. ■

Teniendo en cuenta esta equivalencia, mostramos el siguiente resultado

Teorema 3.2.1 (Principio Débil de Pontryagin-Lagrange (PdPL)) *Para el problema $\mathcal{P}_T(\hat{x}_0)$, asumamos las hipótesis $(H1)$ y $(H2)$. Sea la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso asociado $(\hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima** *tal que $\hat{z} = (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1}) \in \operatorname{int} Z$.*

Entonces existen T vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, tales que

- *Para $t = 1, \cdots, T-1$,*

$$\hat{\psi}_{t-1} = \nabla_x H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \qquad (EA)$$

  con condición final $\hat{\psi}_{T-1} = \nabla_x f_T^0(\hat{x}_T)$.

- *Para cada $t = 0, \cdots, T-1$,*

$$\nabla_u H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) = 0. \qquad (PdPL)_t$$

Por lo tanto u_t es un **punto crítico** *del problema*

$$\max_{u \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t),$$

donde $H_t(x_t, u_t, \psi_t) = f_t^0(x_t, u_t) + \langle \psi_t \ ; \ f_t(x_t, u_t) \rangle$.

Prueba. En vista al Lema 3.1.1, $\hat{z} = (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ es solución del problema estático

$$\min_{z \in \Omega} (-F^0(z)).$$

Debido a las hipótesis $(H1)$ y $(H2)$, y al hecho que $\hat{z} \in \operatorname{int}$ Z, existe un entorno $V(\hat{z}) \subset Z$, en el que F^0 es diferenciable y las F_k^i son de clase C^1. De otro lado, usando el Lema 3.2.1 para $\hat{z} \in Z$, el conjunto

$$\{\nabla_z F_k^i(\hat{z})\}_{\substack{i=1,\cdots,n \\ k=0,\cdots,T-1}} \subseteq \mathbb{R}^{(n+m)T}$$

es linealmente independiente. Por lo tanto $\hat{z}$ es un "punto regular", así que podemos aplicar el Teorema de Lagrange en $\hat{z}$, pues es óptimo. (ver [15, p.47]) Existe entonces un conjunto de multiplicadores $\{\bar{\lambda}_k^i\}_{\substack{i=1,\cdots,n \\ k=0,\cdots,T-1}} \subset \mathbb{R}$, con los cuales definimos $\bar{\lambda} := (\bar{\lambda}_0, \cdots, \bar{\lambda}_{T-1})$, con $\bar{\lambda}_k := (\bar{\lambda}_k^1, \cdots, \bar{\lambda}_k^n) \in \mathbb{R}^n$, tales que

$$\nabla_z L(\hat{z}, \bar{\lambda}) = 0 \in \mathbb{R}^{(n+m)T}, \tag{1}$$

donde la función $L : Z \times \mathbb{R}^{nT} \to \mathbb{R}$, llamado el Lagrangiano, es definido por

$$L(z, \lambda) = -F^0(z) + \sum_{k=0}^{T-1} \sum_{i=1}^{n} \lambda_k^i F_k^i(z).$$

Expresando (1) en sus "coordenadas" x_t y u_t (pues $z = (x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$) tendríamos:

- Con respecto al estado,

$$\begin{array}{rcl} -\nabla_{x_1} F^0(\hat{z}) + \sum \bar{\lambda}_k^i \nabla_{x_1} F_k^i(\hat{z}) & = & 0, \\ -\nabla_{x_2} F^0(\hat{z}) + \sum \bar{\lambda}_k^i \nabla_{x_2} F_k^i(\hat{z}) & = & 0, \\ \cdots \quad \cdots \quad \cdots & & \\ -\nabla_{x_T} F^0(\hat{z}) + \sum \bar{\lambda}_k^i \nabla_{x_T} F_k^i(\hat{z}) & = & 0. \end{array} \tag{Θ}$$

- Con respecto al control,

$$\begin{array}{rcl} -\nabla_{u_0} F^0(\hat{z}) + \sum \bar{\lambda}_k^i \nabla_{u_0} F_k^i(\hat{z}) & = & 0, \\ -\nabla_{u_1} F^0(\hat{z}) + \sum \bar{\lambda}_k^i \nabla_{u_1} F_k^i(\hat{z}) & = & 0, \\ \cdots \quad \cdots \quad \cdots & & \\ -\nabla_{u_{T-1}} F^0(\hat{z}) + \sum \bar{\lambda}_k^i \nabla_{u_{T-1}} F_k^i(\hat{z}) & = & 0. \end{array} \tag{Φ}$$

Aquí las sumatorias están definidas para $i = 1, \cdots, n$ y $k = 0, \cdots, T-1$.

Ahora mostraremos que: i) el sistema (Θ) es equivalente a nuestro sistema adjunto (EA), por lo cual los multiplicadores se pueden interpretar como las variables adjuntas de $\hat{z}$ y, ii) del sistema (Φ) se deduce las igualdades $(PdPL)_t$.

1. El sistema (Θ) es equivalente al Sistema Adjunto (EA) Hagamos algunos cálculos,

$$-\nabla_{x_r} \left(\textstyle\sum_{t=0}^{T-1} f_t^0(x_t, u_t) + f_T^0(x_T) \right) = -\nabla_{x_r} f_r^0(x_r, u_r), \qquad \text{si } r = 1, \cdots, T-1.$$

$$-\nabla_{x_r} \left(\textstyle\sum_{t=0}^{T-1} f_t^0(x_t, u_t) + f_T^0(x_T) \right) = -\nabla_{x_T} f_T^0(x_T), \qquad \text{si } r = T.$$

Con lo cual para $\hat{z}$, tenemos

$$\begin{array}{ll} -\nabla_{x_r} F^0(\hat{z}) = -\nabla_{x_r} f_r^0(\hat{x}_r, \hat{u}_r), & \text{si } r = 1, \cdots, T-1. \\ -\nabla_{x_r} F^0(\hat{z}) = -\nabla_{x_T} f_T^0(\hat{x}_T), & \text{si } r = T. \end{array}$$

Recordando los cálculos del Lema 3.2.1 (ítem 1, cálculo de los gradientes), tenemos para $r = 1, \cdots, T-1$,

$$\begin{array}{rl} \nabla_{x_r} F_k^i(z) = 0, & \text{si } k \in \{0, \cdots, T-1\} \setminus \{r-1, r\}. \\ \nabla_{x_r} F_k^i(z) = e_i, & \text{si } k = r-1. \\ \nabla_{x_r} F_k^i(z) = -\nabla_{x_r} f_r^i(x_r, u_r), & \text{si } k = r. \end{array}$$

Con estos cálculos simplificamos las sumatorias

$$\sum_{k=0}^{T-1} \sum_{i=1}^{n} \bar{\lambda}_k^i \nabla_{x_r} F_k^i(z) = \sum_{i=1}^{n} \bar{\lambda}_{r-1}^i (e_i) - \sum_{i=1}^{n} \bar{\lambda}_r^i (\nabla_{x_r} f_r^i(x_r, u_r)).$$

Similarmente para $r = T$,

$$\begin{array}{rl} \nabla_{x_r} F_k^i(z) = 0, & \text{si } k \in \{0, \cdots, T-2\}. \\ \nabla_{x_r} F_k^i(z) = e_i, & \text{si } k = T-1. \end{array}$$

Así,

$$\sum_{k=0}^{T-1} \sum_{i=1}^{n} \bar{\lambda}_k^i \nabla_{x_T} F_k^i(z) = \sum_{i=1}^{n} \bar{\lambda}_{T-1}^i (e_i).$$

Usando estos cálculos para $\hat{z}$ y reemplazándolos en el sistema (Θ) obtenemos

$$\begin{array}{rl} -\nabla_{x_r} f_r^0(\hat{x}_r, \hat{u}_r) + \bar{\lambda}_{r-1} - \bar{\lambda}_r \circ D_{x_r} f_1(\hat{x}_r, \hat{u}_r) = 0, & \text{si } r = 1, \cdots, T-1, \\ -\nabla_{x_T}(f_T^0(\bar{x}_T)) + \bar{\lambda}_{T-1} = 0, & \text{si } r = T. \end{array}$$

donde hemos definido $\bar{\lambda}_r := (\bar{\lambda}_r^1, \ldots, \bar{\lambda}_r^n) \in \mathbb{R}^n$, para $r = 0, \cdots, T-1$. Usando el hamiltoniano $H_t(x_t, u_t, \bar{\lambda}_t) = f_t^0(x_t, u_t) + \langle \bar{\lambda}_t, f_t(x_t, u_t) \rangle$, vemos

$$\nabla_{x_r} H_r(\hat{x}_r, \hat{u}_r, \bar{\lambda}_r) = \nabla_{x_r} f_r^0(\hat{x}_r, \hat{u}_r) + \bar{\lambda}_r \circ D_{x_r} f_1(\hat{x}_r, \hat{u}_r).$$

Por ello podemos reescribir el sistema anterior como

$$\begin{array}{ccl} \bar{\lambda}_0 & = & \nabla_{x_1} H_1(\hat{x}_1, \hat{u}_1, \bar{\lambda}_1), \\ \bar{\lambda}_1 & = & \nabla_{x_2} H_2(\hat{x}_1, \hat{u}_2, \bar{\lambda}_2), \\ \cdots & \cdots & \cdots \quad \cdots \quad \cdots \quad \cdots \\ \bar{\lambda}_{T-2} & = & \nabla_{x_{T-1}} H_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1}, \bar{\lambda}_{T-1}), \\ \bar{\lambda}_{T-1} & = & \nabla_{x_T}(f_T^0(\hat{x}_T)). \end{array}$$

Tomando para $r = 0, \cdots, T-1$,

$$\hat{\psi}_r := \bar{\lambda}_r \in \mathbb{R}^n,$$

obtenemos el "Sistema Adjunto" (EA). Por ello los vectores $\hat{\psi}_0, \cdots \hat{\psi}_{T-1}$, serán las variables adjuntas asociadas a la sucesiones de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$.

2. El sistema (Φ) es equivalente a $(PdPL)_t$ Similar a los cálculos anteriores obtenemos, para $r = 0, \cdots, T-1$,

$$-\nabla_{u_r} F(z) = -\nabla_{u_r}(\sum_{t=0}^{T-1} f_t^0(x_t, u_t) + f_T^0(x_T)) = -\nabla_{u_r} f_r^0(x_r, u_r).$$

También,

$$\begin{aligned} \nabla_{u_r} F_k^i(z) &= 0, && \text{si } r \neq k, \\ \nabla_{u_r} F_k^i(z) &= -\nabla_{u_r} f_r^i(x_r, u_r), && \text{si } r = k. \end{aligned}$$

Por lo cual,

$$\sum_{k=0}^{T-1} \sum_{i=1}^{n} \bar{\lambda}_k^i \nabla_{u_r} F_k^i(z) = \sum_{i=1}^{n} \bar{\lambda}_r^i \nabla_{u_r} F_r^i(z) = -\bar{\lambda}_r \circ D_u f_r(x_r, u_r).$$

Por ello tenemos en $\hat{z}$, para $r = 1, \cdots, T$,

$$-\nabla_{u_r} F(\hat{z}) + \sum_{k=0}^{T-1} \sum_{i=1}^{n} \bar{\lambda}_k^i \nabla_{u_r} F_k^i(\hat{z}) = -\nabla_{u_r} f_r^0(\hat{x}_r, \hat{u}_r) - \bar{\lambda}_r \circ D_u f_r(\hat{x}_r, \hat{u}_r).$$

Haciendo uso del sistema (Φ), deducímos que para $r = 0, \cdots, T-1$,

$$0 = -\nabla_{u_r} f_r^0(\hat{x}_r, \hat{u}_r) - \bar{\lambda}_r \circ D_u f_r(\hat{x}_r, \hat{u}_r).$$

Debido a que $\hat{\psi}_r := \bar{\lambda}_r \in \mathbb{R}^n$ obtenemos,

$$0 = \nabla_{u_r} f_r^0(\hat{x}_r, \hat{u}_r) + \hat{\psi}_r \circ D_u f_r(\hat{x}_r, \hat{u}_r).$$

Nuevamente usando el hamiltoniano $H_t(x_t, u_t, \bar{\lambda}_t) = f_t^0(x_t, u_t) + \langle \bar{\lambda}_t, f_t(x_t, u_t) \rangle$, vemos que

$$\nabla_{u_r} H_r(\hat{x}_r, \hat{u}_r, \bar{\lambda}_r) = \nabla_{u_r} f_r^0(\hat{x}_r, \hat{u}_r) + \bar{\lambda}_r \circ D_{u_r} f_1(\hat{x}_r, \hat{u}_r),$$

por lo que obtenemos

$$0 = \nabla_{u_r} H_r(\hat{x}_r, \hat{u}_r, \hat{\psi}_r).$$

Interpretando este resultado para el problema

$$\max_{u \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t),$$

deducimos que $\hat{u}_t$ es un punto crítico, con lo cual terminamos esta demostración. ■

Una prueba enteramente análoga se tiene para el problema con extremo derecho fijo $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$. Recordemos su definición

$$\begin{array}{ll} \max & B_T(\hat{x}_0, u_0, \cdots, u_{T-1}) := \sum_{t=0}^{T-1} f_t^0(x_t, u_t) + f_T^0(\hat{x}_T) \\ \text{s.a.} & \text{para todo } t = 0, \cdots, T-1,\ u_t \in U_t, \\ & x_{t+1} \in X_{t+1},\ \ x_{t+1} = f_t(x_t, u_t), \\ & x_0 = \hat{x}_0 \in X_0 \ \text{ y } \ x_T = \hat{x}_T. \end{array} \qquad (\mathcal{P}_T(\hat{x}_0, \hat{x}_T))$$

Corolario 3.2.1 (Extremo Fijo) *Para el problema $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$, asumamos las hipótesis $(H1)$ y $(H2)$. Sea la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso asociado $(\hat{x}_0, \cdots, \hat{x}_T)$, una* **solución** *tal que $\hat{z} = (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1}) \in \operatorname{int} Z$.*

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, tales que

- *Para $t = 1, \cdots, N-1$*

$$\hat{\psi}_{t-1} = \nabla_x H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t). \qquad (EA)$$

- *Para cada $t = 0, \cdots, N-1$,*

$$\nabla_u H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) = 0. \qquad (PdPL)_t$$

Prueba. La prueba se obtiene realizando la misma transformación del problema $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$, en problema estático, con la única diferencia que aquí, ya no se tendrá en cuenta la última variable (x_T) es por ello que desaparece la condición final, en el sistema de ecuaciones (EA). ■

Observación. Digamos algo sobre las hipótesis. Ellas son generales (se exige que se cumpla (H1) y (H2) para todo el dominio de f_t^0 y f_t), pero en realidad bastaría pedir que se satisfagan en un entorno del punto en cuestión (al que aplicamos la condición necesaria). Sin embargo, hemos procedido de esa manera pues en los problemas prácticos por lo general no se tiene la menor idea de cual puede ser un candidato, al que aplicaremos el teorema, por ello es que al cumplirse las hipótesis de modo general, podemos en principio "rastrear" al óptimo como uno de los puntos que cumplen nuestra condición.

Una vez probado el resultado anterior, naturalmente surge la idea de usar también el teorema de la condición suficiente de segundo orden. En esa dirección tenemos un resultado que nos permite tener una cierta metodología para comprobar si al menos localmente un punto es óptimo. Hemos realizado los cálculos necesarios para obtener dicha condición. Hemos de decir que lamentablemente, este resultado no es tan elegante y suscinto como el principio débil de Pontryagin-Lagrange, sino que exige demasiados cálculos, por lo que su aplicación práctica es inviable.

Extensiones El teorema de Pontryagin discreto, tal y como esta formulado, es una condición necesaria para puntos que satisfacen la condición (I), la cual es equivalente a que $x_t \in \operatorname{int} \mathrm{X_t}$, $u_s \in \operatorname{int} \mathrm{U_s}$, para cada $t = 1, \cdots, T$ y $s = 0, \cdots, T-1$. Es natural preguntar ¿Qué condiciones habría para puntos que no satisfacen esta condición? Es decir si es que para algún t o s, se tiene que $x_t \in \partial X_t$ ó $u_s \in \partial U_s$ (puntos de la frontera). En este caso podemos considerar a estos conjuntos también como restricciones adicionales al problema. De modo que podamos usar condiciones necesarias para casos más generales. Con esta finalidad hemos de exigir que las funciones f_t y f_t^0 estén definidas sobre un abierto W_t que contenga a $X_t \times U_t$. Con lo cual podremos considerar a estos conjuntos como restricciones para nuestro problema. En base a estas ideas trabajaremos para los siguientes "Principios Débiles de Pontryagin".

Notemos que a pesar de que las funciones estén definidas en un conjunto más amplio W_t, el problema $\mathcal{P}_T(\hat{x}_0)$ continua siendo el mismo. Es decir nos sigue interesando sólo los caminos que evolucionan en los conjuntos X_t usando los controles U_t.

En los libros que hemos consultado no se hace referencia a estas cuestiones pues por lo general toman las funciones f_t y f_t^0 definidas en todo el espacio ($W_t = \mathbb{R}^n \times \mathbb{R}^m$). Sin embargo a nosotros no nos parece natural ese supuesto tácito, pues suponemos que en muchos problemas reales no tendría sentido definir las funciones de beneficios en todo el espacio, entendiendo que en varias ocasiones los controles son limitados (acotados, máxima fuerza, mínima temperatura, etc.), así como los estados posibles del sistema. Por todo esto nos parece justo hacer todas estas observaciones.

Pues bien, en este nuevo marco (cuando las funciones f_t y f_t^0 están definidas en abiertos) es que trabajaremos para poder dar condiciones de optimalidad para casos más generales que el interior, por ejemplo cuando $\bar{x}_t \in \partial X_t$ ó $\bar{u}_t \in \partial U_t$ (frontera).

3.3 Principio Débil: Fritz John

Como decíamos anteriormente para conseguir condiciones de optimalidad más generales necesitamos que nuestras funciones f_t y f_t^0, estén definidas en dominios más amplios. Supongamos que para $t = 0, 1, \cdots, T-1$, estas funciones están definidas en el dominio $W_t = A_t \times B_t$, con $A_t \subset \mathbb{R}^n$ y $B_t \subset \mathbb{R}^m$ ambos son **conjuntos abiertos** que satisfazen $X_t \subset A_t$ y $U_t \subset B_t$. También suponemos que f_T^0 está definida sobre un abierto $W_T \subset \mathbb{R}^n$ con $X_T \subset W_T$.

En las próximas secciones de este capítulo **siempre consideraremos que las funciones f_t y f_t^0 están definidos en entornos abiertos que contienen a** $X_t \times U_t$.

En esta sección demostraremos una condición necesaria de optimalidad para el problema $\mathcal{P}_T(\hat{x}_0)$, cuando satisface la siguiente hipótesis (R).

Para $t = 1, \cdots, T$,

$$X_t = \{x \in A_t : p_t(x) \leq 0\}, \tag{R}$$

donde $p_t : A_t \to \mathbb{R}^{n_t}$, es una función diferenciable, y para $t = 0, \cdots, T-1$,

$$U_t = \{x \in B_t : q_t(u) \leq 0\}, \tag{R}$$

con $q_t : A_t \to \mathbb{R}^{m_t}$, función diferenciable.

Siguiendo las ideas anteriores tenemos que transformar el problema $\mathcal{P}_T(\hat{x}_0)$ (en este nuevo contexto) a un problema de optimización estática.

Lema 3.3.1 (Transformación: Dinámico a Estático) *En el contexto presente el problema de control $\mathcal{P}_T(\hat{x}_0)$ es equivalente a un problema de optimización estático.*

Prueba. Denotemos $W = \prod_{s=1}^{T} A_s \times \prod_{t=0}^{T-1} B_t \subset \mathbb{R}^{(n+m)T}$ y a sus elementos como $w := (x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$.

Definamos $F^0 : W \to \mathbb{R}$, por

$$F^0(z) = \sum_{k=0}^{T-1} f_k^0(x_k, u_k) + f_T^0(x_T),$$

y

$$\Omega = \left\{ w \in W \;\middle|\; \begin{array}{l} \text{para } \; i = 1, \cdots, n; \; t = 0, \cdots, T-1; \; s = 1, \cdots \\ \qquad\quad \cdots, T; \; j_s = 1, \cdots, n_s \; \text{ y } i_t = 1, \cdots, m_t, \\ F_t^i(w) = 0, \; P_s^{j_s}(w) \leq 0, \; Q_t^{i_t}(w) \leq 0. \end{array} \right\},$$

donde, para $t = 0, \cdots, T-1,$; $i = 1, \cdots, n;$ $s = 1, \cdots, T;$ $j_s = 1, \cdots, n_s$ y $i_t = 1, \cdots, m_t$,

$$\begin{aligned} F_t^i(w) &= x_{t+1}^i - f_t^i(x_t, u_t), \\ P_s^{j_s}(w) &= p_s^{j_s}(x_s), \\ Q_t^{i_t}(w) &= q_t^{i_t}(u_t). \end{aligned}$$

Con estas definiciones el problema de control discreto $\mathcal{P}_T(\hat{x}_0)$ es equivalente al problema de optimización estática con restricciones de igualdades y desigualdades,

$$\max_{w \in \Omega} F^0(w) \quad \equiv \quad \max_{\substack{F_k^i(w)=0, \\ P_s^{j_s}(w) \leq 0, \\ Q_t^{i_t}(w) \leq 0.}} F^0(w).$$

De este se deduce la equivalencia. ∎

En lo que sigue de esta subsección haremos uso de las siguientes hipótesis:

1. Para $t = 0, \cdots, T-1$,

$$f_t^0 \text{ es diferenciable en } A_t \times B_t. \qquad (H1)_J$$

2. Para $t = 0, \cdots, T-1$,

$$f_t \text{ es de clase } C^1 \text{ en } A_t \times B_t. \qquad (H2)_J$$

3. Para $s = 1, \cdots, T$,

$$p_s \text{ es diferenciable en } A_s. \qquad (H3)_J$$

4. Para $t = 0, \cdots, T-1$,

$$q_t \text{ es diferenciable en } B_t. \qquad (H4)_J$$

En este nuevo contexto, los **hamiltonianos** asociados al problema $\mathcal{P}_T(\hat{x}_0)$ toma una forma más general. Para $t = 0, \cdots, T-1$, el hamiltoniano $H_t : A_t \times B_t \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n_t} \to \mathbb{R}$ se define como

$$H_t(x, u, a_0, \psi, \nu) = a_0 f_t^0(x, u) + \langle \psi, f_t(x, u) \rangle - \langle \nu, p_t(x_t) \rangle.$$

En este nuevo marco conceptual, con estas hipótesis y definiciones podemos probar el siguiente teorema.

Teorema 3.3.1 (Principio Débil de Pontryagin-John Discreto (PdPJ)) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ asumamos $(H1)_J - (H4)_J$. Sea $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ una sucesión de decisiones, con su proceso asociado $(\hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima**.

Entonces existen $a_0 \geq 0$, $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, $\{\hat{\nu}_s \in \mathbb{R}_+^{n_s} : s = 1, \cdots, T\}$, $\{\hat{\mu}_t \in \mathbb{R}_+^{m_t} : t = 0, \cdots, T-1\}$ **no todos nulos**, *tales que:*

1. *Para $t = 1, \cdots, T-1$,*

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t, \hat{\nu}_t), \qquad (EA)_J$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T} \left(a_0 f_T^0(\hat{x}_T) - \langle \hat{\nu}_T, p_T(\hat{x}_T) \rangle \right)$.

2. *Para $t = 1, \cdots T$ y $j_t = 1, \cdots, n_t$,*

$$\hat{\nu}_t^{j_t} p_t^{j_t}(\hat{x}_t) = 0. \qquad (C(x))_J$$

(Condición de complementariedad para x_t).

3. *Para $t = 0, \cdots, T-1$,*

$$\nabla_{u_t} \Big(H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t, \hat{\nu}_t) - \langle \hat{\mu}_t, q_t(\hat{u}_t) \rangle \Big) = 0. \qquad (PdPJ)$$

4. *Para* $t = 0, \cdots T-1$ *y* $i_t = 1, \cdots, m_t$,

$$\hat{\mu}_t^{i_t} q_t^{i_t}(\hat{u}_t) = 0. \qquad (C(u))_J$$

(Condición de complementariedad para u_t*).*

Por lo tanto, de los items 3 y 4, para $t = 0, \cdots, T-1$, $\hat{u}_t$ *es un* **punto crítico** *del problema*

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, a_0, \hat{\psi}_t, \hat{\nu}_t), \qquad (PMdPJ)_t$$

y de los items 1 y 2, los vectores $\hat{\psi}_t$ *son las variables adjuntas.*

Prueba. La prueba consiste en aplicar la condición necesaria de optimalidad de Fritz John al problema,

$$\max_{\substack{F_k^i(w)=0,\\ P_s^{j_s}(w)\le 0,\\ Q_t^{i_t}(w)\le 0.}} F^0(w) \qquad \equiv \qquad \min_{\substack{F_k^i(w)=0,\\ P_s^{j_s}(w)\le 0,\\ Q_t^{i_t}(w)\le 0.}} -F^0(w),$$

que es equivalente a nuestro problema de control óptimo discreto $\mathcal{P}_T(\hat{x}_0)$, por el Lema 3.3.1. De ello $\hat{w} := (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1}) \in W$ es una de sus soluciones óptimas.

Apliquemos el teorema de Fritz John (ver [15, p.206], [1, p.182]). Definamos nuestro Lagrangiano,

$$L : W \times \mathbb{R} \times \mathbb{R}^{nT} \times \prod_{s=1}^{T} \mathbb{R}^{n_s} \times \prod_{t=0}^{T-1} \mathbb{R}^{m_t} \to \mathbb{R},$$

$$L(w, a, \psi_k^i, \nu_s^{j_s}, \mu_t^{i_t}) = -aF^0(w) + \textstyle\sum_{k=0}^{T-1}\sum_{i=1}^{n} \psi_k^i F_k^i(w) +$$

$$+ \textstyle\sum_{s=1}^{T}\sum_{j_s=1}^{n_s} \nu_s^{i_s} P_s^{i_s}(w) + \sum_{t=0}^{T-1}\sum_{i_t=1}^{m_t} \mu_k^i Q_t^{i_t}(w).$$

Entonces aplicando el teorema de Fritz para $\hat{z}$, existen números reales **no todos nulos**,

$$a_0 \ge 0, \quad (\bar{\psi}_k^i)_{\substack{i=1,\cdots,n\\ k=0,\cdots,T-1}}, \quad (\bar{\nu}_s^{j_s})_{\substack{j_s=1,\cdots,n_s\\ s=1,\cdots,T}} \quad \text{y} \quad (\bar{\mu}_t^{i_t})_{\substack{i_t=1,\cdots,m_t,\\ t=0,\cdots,T-1}}$$

tales que para $j_s = 1, \cdots, n_s$, $s = 1, \cdots, T$, $i_t = 1, \cdots, m_t$, y $t = 0, \cdots, T-1$,

1.

$$\nabla_w L(\hat{w}, a_0, \bar{\psi}_k^i, \bar{\nu}_s^{j_s}, \bar{\mu}_t^{i_t}) = 0, \qquad (Nul)$$

2.

$$\bar{\nu}_s^{j_s} P_s^{j_s}(\hat{w}) = 0 \quad \text{y} \quad \bar{\mu}_t^{i_t} Q_t^{i_t}(\hat{w}) = 0, \qquad (Com)$$

3.

$$\bar{\nu}_s^{j_s} \geq 0 \quad \text{y} \quad \bar{\mu}_t^{i_t} \geq 0. \qquad (Pos)$$

El gradiente en (Nul) es

$$-a_0 \nabla_w F^0(\hat{w}) + \sum_{k=0}^{T-1} \sum_{i=1}^{n} \bar{\psi}_k^i \nabla_w F_k^i(\hat{w}) + \sum_{s=1}^{T} \sum_{j_s=1}^{n_s} \bar{\nu}_s^{i_s} \nabla_w P_s^{i_s}(\hat{w}) + \sum_{t=0}^{T-1} \sum_{i_t=1}^{m_t} \bar{\mu}_k^i \nabla_w Q_t^{i_t}(\hat{w}).$$

Aquí, para $w = (x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$,

$$\nabla_w F_k^i(w) = (\nabla_{x_1} F_k^i(z), , \cdots, \nabla_{x_T} F_k^i(z); \nabla_{u_0} F_k^i(z), \cdots, \nabla_{u_{T-1}} F_k^i(z)),$$

$$\nabla_w P_s^{j_s}(w) = (\nabla_{x_1} P_s^{j_s}(w), \cdots, \nabla_{x_T} P_s^{j_s}(w); \nabla_{u_0} P_s^{j_s}(w), \cdots, \nabla_{u_{T-1}} P_s^{j_s}(w)),$$

y

$$\nabla_w Q_t^{i_t}(w) = (\nabla_{x_1} Q_t^{i_t}(w), \cdots, \nabla_{x_T} Q_t^{i_t}(w); \nabla_{u_0} Q_t^{i_t}(w), \cdots, \nabla_{u_{T-1}} Q_t^{i_t}(w)).$$

Reemplazándolo en (Nul) tendríamos dos sistemas de ecuaciones:

Sistema Θ: Para $l = 1, \cdots, T$,

$$-a_0 \nabla_{x_l} F^0(\hat{w}) + \sum \bar{\psi}_k^i \nabla_{x_l} F_k^i(\hat{w}) + \sum \bar{\nu}_s^{i_s} \nabla_{x_l} P_s^{i_s}(\hat{w}) + \sum \bar{\mu}_k^i \nabla_{x_l} Q_t^{i_t}(\hat{w}) = 0,$$

donde

$$\begin{aligned} \textstyle\sum \bar{\psi}_k^i \nabla_{x_l} F_k^i(\hat{w}) &= \textstyle\sum_{k=0}^{T-1} \sum_{i=1}^{n} \bar{\psi}_k^i \nabla_{x_l} F_k^i(\hat{w}), \\ \textstyle\sum \bar{\nu}_s^{i_s} \nabla_{x_l} P_s^{i_s}(\hat{w}) &= \textstyle\sum_{s=1}^{T} \sum_{j_s=1}^{n_s} \bar{\nu}_s^{i_s} \nabla_{x_l} P_s^{i_s}(\hat{w}), \\ \textstyle\sum \bar{\mu}_k^i \nabla_{x_l} Q_t^{i_t}(\hat{w}) &= \textstyle\sum_{t=0}^{T-1} \sum_{i_t=1}^{m_t} \bar{\mu}_k^i \nabla_{x_l} Q_t^{i_t}(\hat{w}). \end{aligned}$$

Sistema Φ: Para $r = 0, \cdots, T-1$,

$$-a_0 \nabla_{u_r} F^0(\hat{w}) + \sum \bar{\psi}_k^i \nabla_{u_r} F_k^i(\hat{w}) + \sum \bar{\nu}_s^{i_s} \nabla_{u_r} P_s^{i_s}(\hat{w}) + \sum \bar{\mu}_k^i \nabla_{u_r} Q_t^{i_t}(\hat{w}) = 0,$$

donde

$$\begin{aligned} \textstyle\sum \bar{\psi}_k^i \nabla_{u_r} F_k^i(\hat{w}) &= \textstyle\sum_{k=0}^{T-1} \sum_{i=1}^{n} \bar{\psi}_k^i \nabla_{u_r} F_k^i(\hat{w}), \\ \textstyle\sum \bar{\nu}_s^{i_s} \nabla_{u_r} P_s^{i_s}(\hat{w}) &= \textstyle\sum_{s=1}^{T} \sum_{j_s=1}^{n_s} \bar{\nu}_s^{i_s} \nabla_{u_r} P_s^{i_s}(\hat{w}), \\ \textstyle\sum \bar{\mu}_k^i \nabla_{u_r} Q_t^{i_t}(\hat{w}) &= \textstyle\sum_{t=0}^{T-1} \sum_{i_t=1}^{m_t} \bar{\mu}_k^i \nabla_{u_r} Q_t^{i_t}(\hat{w}). \end{aligned}$$

Similarmente a la demostración del Teorema 3.2.1 usaremos ambos sistemas para deducir el teorema. Del sistema Θ deduciremos el sistema de ecuaciones $(EA)_J$ y desde el sistema Φ deduciremos $(PdPJ)$. Finalmente usando la condición de complementariedad (Com) deduciremos $(C(x))_J$ y $(C(u))_J$.

1. Sistema $\Theta \equiv$ Sistema Adjunto $(EA)_J$ Hagamos algunos cálculos. Para $l = 1, \cdots, T-1$, tenemos

$$-a_0 \nabla_{x_l} F^0(\hat{w}) = -a_0 \nabla_{x_l} (\sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t) + f_T^0(\hat{x}_T)) = -a_0 \nabla_{x_l} f_l^0(\hat{x}_l, \hat{u}_l),$$

y para $l = T$,

$$-a_0 \nabla_{x_l} F^0(\hat{w}) = -a_0 \nabla_{x_l} (\sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t) + f_T^0(\hat{x}_T)) = -a_0 \nabla_{x_T} f_T^0(\hat{x}_T).$$

Las funciones F_k^i son las mismas con las que se trabajaron en el Teorema 3.2.1, por ello los cálculos de los gradientes hechos en el Lema 3.2.1 son también válidos para este teorema. Dado $w = (x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$, tenemos:

Para $r = 1, \cdots, T-1$,

$$\begin{array}{ll} \nabla_{x_r} F_k^i(w) = 0, & \text{si } k \in \{0, \cdots, T-1\} \setminus \{r-1, r\}, \\ \nabla_{x_r} F_k^i(w) = e_i, & \text{si } k = r-1, \\ \nabla_{x_r} F_k^i(w) = -\nabla_{x_r} f_r^i(x_r, u_r), & \text{si } k = r. \end{array}$$

De donde,

$$\sum_{k=0}^{T-1} \sum_{i=1}^{n} \psi_k^i \nabla_{x_r} F_k^i(w) = \sum_{i=1}^{n} \psi_{r-1}^i(e_i) - \sum_{i=1}^{n} \psi_r^i(\nabla_{x_r} f_r^i(x_r, u_r)).$$

Similarmente, para $r = T$,

$$\begin{array}{ll} \nabla_{x_r} F_k^i(w) = 0 & \text{si } k \in \{0, \cdots, T-2\}, \\ \nabla_{x_r} F_k^i(w) = e_i & \text{si } k = T-1. \end{array}$$

De donde,

$$\sum_{k=0}^{T-1} \sum_{i=1}^{n} \psi_k^i \nabla_{x_T} F_k^i(w) = \sum_{i=1}^{n} \psi_{T-1}^i(e_i).$$

Calculando también los gradientes de las funciones $P_s^{j_s}$ y $Q_t^{i_t}$, obtenemos para $l = 1, \cdots, T$,

$$\begin{array}{ll} \nabla_{x_l} P_s^{j_s}(w) = 0, & \text{si } s \neq l, \\ \nabla_{x_l} P_s^{j_s}(w) = \nabla_{x_l} p_l^{j_l}(x_l), & \text{si } s = l. \end{array}$$

Y para $r = 0, \cdots, T-1$,

$$\begin{array}{ll} \nabla_{u_s} Q_t^{i_t}(w) = 0, & \text{si } s \neq t, \\ \nabla_{u_s} Q_t^{i_t}(w) = \nabla_{u_s} q_t^{i_t}(u_t), & \text{si } s = t. \end{array}$$

Juntando todos estos cálculos y reemplazándolos en el sistema Θ tenemos para $l = 1, \cdots, T-1$,

$$-a_0 \nabla_{x_l} f_l^0(\hat{x}_l, \hat{u}_l) + \bar{\psi}_{l-1} - \bar{\psi}_l \circ D_{x_l} f_l(\hat{x}_l, \hat{u}_l) + \bar{\nu}_l \circ D_{x_l} p_l(\hat{x}_l) = 0,$$

y para $l = T$,

$$-a_0 \nabla_{x_T} f_T^0(\hat{x}_T) + \bar{\psi}_{T-1} + \bar{\nu}_T \circ D_{x_T} p_T(\hat{x}_T) = 0,$$

donde, para $r = 0, \cdots, T-1$ y $s = 1, \cdots, T$,

$$\bar{\psi}_r = (\bar{\psi}_r^1, \cdots, \bar{\psi}_r^n) \in \mathbb{R}^n \quad \text{y} \quad \bar{\nu}_s = (\bar{\nu}_s^1, \cdots, \bar{\nu}_s^{n_s}) \in \mathbb{R}^{n_s}.$$

Usando el Hamiltoniano definido después del Lema 3.3.1, tenemos

$$\nabla_{x_t} H_t(x_t, u_t, a_0, \psi_t, \nu_t) = a_0 \nabla_{x_t} f_t^0(x_t, u_t) + \psi_t \circ D_{x_t} f_t(x_t, u_t) - \nu_t \circ D_{x_t} p_t(x_t).$$

Por lo tanto el sistema Θ puede ser escrito como,

$$\begin{array}{lcl}
\bar{\psi}_0 & = & \nabla_{x_1} H_1(\hat{x}_1, \hat{u}_1, a_0, \bar{\psi}_1, \bar{\nu}_1), \\
\bar{\psi}_1 & = & \nabla_{x_2} H_2(\hat{x}_2, \hat{u}_2, a_0, \bar{\psi}_2, \bar{\nu}_2), \\
\cdots & \cdots & \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
\bar{\psi}_{T-2} & = & \nabla_{x_{T-1}} H_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1}, a_0, \bar{\psi}_{T-1}, \bar{\nu}_{T-1}), \\
\bar{\psi}_{T-1} & = & a_0 \nabla_{x_T}(f_T^0(\hat{x}_T)) - \nu_T \circ D_{x_T} p_{x_T}(\hat{x}_T).
\end{array}$$

Tomando, para $t = 0, \cdots, T-1$, $\hat{\psi}(t) := \bar{\psi}_t$ y $\hat{\nu}(t) := \bar{\nu}_t$, y considerando la condición (Pos), tenemos $\hat{\nu}_s \in \mathbb{R}_+^{n_s}$. Con esto vemos que lo anterior es equivalente al sistema adjunto de nuestro teorema. Por lo tanto hemos logrado demostrar que se cumple $(EA)_J$.

Por otra parte, de la condición (Com) se deduce,

$$\hat{\nu}_s^{j_s} P_s^{j_s}(w) = \hat{\nu}_s^{j_s} p_s^{j_s}(\hat{x}_s) = 0,$$

para $s = 1, \cdots, T$ y $j_s = 1, \cdots, n_s$. Por lo tanto obtenemos también $(C(x))_J$.

2. Sistema $(\Phi) \equiv$ **(PdPJ)** Similar a los cálculos anteriores obtenemos para $t = 0, \cdots, T-1$,

$$-\nabla_{u_t} F^0(w) = -\nabla_{u_t} f_t^0(x_t, u_t),$$

$$\sum_{k=0}^{T-1} \sum_{i=1}^{n} \bar{\psi}_k^i \nabla_{u_t} F_k^i(w) = \sum_{i=1}^{n} \bar{\psi}_t^i \nabla_{u_t} F_t^i(w) = -\bar{\psi}_t \circ D_u f_t(x_t, u_t).$$

De otra parte para $t = 0, \cdots, T$,

$$\begin{array}{ll}
\nabla_{u_t} P_s^{j_s}(w) = 0 & \text{si } s = 1, \cdots, T, \\
\nabla_{u_t} Q_s^{j_s}(w) = 0 & \text{si } s \neq t, \\
\nabla_{u_t} Q_s^{j_s}(w) = \nabla_{u_t} q_s^{j_s}(u_s) & \text{si } s = t.
\end{array}$$

Reemplazando estos cálculos en el sistema Φ obtenemos, para $t = 1, \cdots, T$,

$$0 = -a_0 \nabla_{u_t} f_t^0(\hat{x}_t, \hat{u}_t) - \bar{\psi}_t \circ D_u f_t(\hat{x}_t, \hat{u}_t) + \sum_{i_t=1}^{m_t} \bar{\mu}_t^{i_t} \nabla_{u_t} q_t^{i_t}(\hat{u}_t).$$

Nuevamente usando el hamiltoniano, calculamos para $t = 0, \cdots, T-1$,

$$\nabla_{u_t} H_t(x_t, u_t, a_0, \psi_t, \nu_t) = a_0 \nabla_{u_t} f_t^0(x_t, u_t) + \psi_t \circ D_{u_t} f_t(x_t, u_t).$$

Con esto deducimos que el sistema Φ es equivalente, para $t = 0, \cdots, T-1$, a

$$0 = -\nabla_{u_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \bar{\psi}_t, \bar{\nu}_t) + \sum_{i_t=1}^{m_t} \bar{\mu}_t^{i_t} \nabla_{u_t} q_t^{i_t}(\hat{u}_t).$$

Nuevamente tomando $\hat{\psi}(t) := \bar{\psi}_t$, $\hat{\nu}(t) := \bar{\nu}_t$ y $\hat{\mu}(t) := \bar{\mu}_t$, para $t = 0, \cdots, T-1$ y considerando la condición (Pos) tenemos $\hat{\mu}_t \in \mathbb{R}_+^{m_t}$. Por ello, usando estos vectores, este sistema será equivalente al sistema $(PdPJ)$ de nuestro teorema.

Finalmente de la condición (Com), tenemos para $t = 0, \cdots T-1$ y $i_t = 1, \cdots, m_t$,

$$\hat{\mu}_t^{i_t} Q_t^{i_t}(w) = \hat{\mu}_t^{i_t} q_t^{i_t}(u_t) = 0,$$

que es precisamente $(C(u))_J$, y por lo tanto la demostración del teorema queda establecida. ■

Observemos que podemos interpretar el resultado anterior del siguiente modo. Para $t = 0, \cdots, T-1$, de la estructura de U_t, se tiene las equivalencias:

$$\max_{u \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t) \quad \equiv \quad \max_{q_t^{i_t}(u_t) \leq 0} H_t(\hat{x}_t, u_t, \hat{\psi}_t). \qquad (PMdPJ)_t$$

En particular la condición necesaria de John para este problema en el punto $\hat{u}_t$ será:

Existen números b_0, $(\hat{\eta}_t^{i_t})_{i_t=1,\cdots,n_t}$ **no todos nulos** tales que:

1.

$$-b_0 \nabla_{u_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t, \bar{\nu}_t) + \sum_{i_r=1}^{m_r} \hat{\eta}_t^{i_t} \nabla_{u_t} q_t^{i_t}(\hat{u}_t) = 0. \qquad (Nul)_t$$

2.

$$\hat{\eta}_t^{i_t} q_t^{i_t}(\hat{u}_t) = 0. \qquad (Com)_t$$

3.

$$\hat{\eta}_t^{i_t} \geq 0. \qquad (Pos)_t$$

Observando el resultado anterior, vemos que si tomamos $\hat{\eta}_t^{i_t} := \hat{\mu}_t^{i_t}$ obtenemos que $\hat{u}_t$ satisface la condición necesaria de Fritz John para el subproblema $(PMdPJ)_t$, y por ello es un punto crítico.

Un caso particular del teorema anterior será cuando $x_t \in \operatorname{int}(X_t)$. Veremos que el teorema toma una forma más aproximada al principio fuerte de pontryaguin, con la ventaja adicional que no exigimos que u_t sea un punto interior.

Corolario 3.3.1 (Estados interiores) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ asumamos las hipótesis $(H1)_J - (H4)_J$. Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso asociado $(\hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima**. *Supongamos además que $\hat{x}_t \in \operatorname{int} X_t$, para $t = 1, \cdots, T$.*

Entonces existen $a_0 \in \mathbb{R}_+$ y vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, $\{\hat{\mu}_t \in \mathbb{R}_+^{m_t} : t = 0, \cdots, T-1\}$ **no todos nulos**, *tales que:*

1. *Para $t = 1, \cdots, T-1$,*

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t), \qquad (EA)_J$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T}\left(a_0 f_T^0(\hat{x}_T)\right)$.

2. *Para $t = 0, \cdots, T-1$,*

$$\nabla_{u_t}\Big(H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t) - \langle \hat{\mu}_t, q_t(\hat{u}_t)\rangle\Big) = 0. \qquad (PdPJ)$$

3. *Para $t = 0, \cdots T-1$ y $i_t = 1, \cdots, m_t$,*

$$\hat{\mu}_t^{i_t} q_t^{i_t}(\hat{u}_t) = 0. \qquad (C(u))_J$$

(Condición de complementariedad para u_t).

Por lo tanto de los ítems 2. y 3. para $t = 0, \cdots, N-1$, $\hat{u}_t$ es punto crítico del subproblema

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, a_0, \hat{\psi}_t),$$

donde $H_t(x, u, a_0, \psi) := a_0 f_t^0(x, u) + \langle \psi, f_t(x, u)\rangle$ es el hamiltoniano.

Prueba. Este corolario se deduce inmediatamente del Teorema 3.3.1, observando que desde $\hat{x}_t \in \operatorname{int} X_t$, se tiene $p_t^{j_t}(\hat{x}_t) < 0$, y por lo tanto desde la complemetariedad $(C(x))_J$ se deduce que $\bar{\nu}_t^{j_t} = 0$, para $t = 1, \cdots, T$ y $j_t = 1, \cdots, n_t$. Por ello los hamiltonianos del teorema anterior se reducen a

$$H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t, \bar{\nu}_t) = H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t).$$

De donde usando el teorema obtenemos el corolario, solamente teniendo en cuenta que $\bar{\nu}_t^{j_t} = 0$ para $t = 1, \cdots, T$ y $j_t = 1, \cdots, n_t$. ■

Vemos que este corolario refleja más fielmente el Principio Fuerte del Máximo (PFP) (ver Proposición 3.1.2).

El siguiente paso que podemos dar es estudiar en que sentido se traducirían las condiciones de calificación (Mangasarian-Fromovitz, etc.) para nuestro problema de control discreto desde su problema equivalente de optimización estático. Como consecuencia de dicho estudio obtendremos un teorema más preciso que el Teorema 3.3.1 (PdPJ), en el que usaremos el Teorema de Karush-Kuhn-Tucker (KKT).

3.4 Principio Débil: Karush Kuhn Tucker

En la presente sección seguimos trabajando con las mismas hipótesis de la sección anterior es decir el problema $\mathcal{P}_T(\hat{x}_0)$ satisface las hipótesis $(H1)_J - (H4)_J$ y para las restricciones la hipótesis (R).

Lo nuevo aquí es el estudio de las condiciones de calificación que deberá cumplir nuestro proceso óptimo visto como "estático" (problema equivalente), para que se satisfaga la condición necesaria de optimalidad de KKT.

Del Lema 3.3.1 sabemos que bajo las hipótesis anteriores nuestro problema $\mathcal{P}_T(\hat{x}_0)$ es equivalente al problema estático $\mathcal{PE}_T(\hat{x}_0)$.

$$\max_{w\in\Omega} F^0(w) \quad \equiv \quad \max_{\substack{F_k^i(w)=0,\\ P_s^{j_s}(w)\le 0,\\ Q_t^{i_t}(w)\le 0.}} F^0(w). \qquad \mathcal{PE}_T(\hat{x}_0)$$

Para este problema usaremos el teorema de KKT antes de ello estudiemos primero las condiciones de calificación.

3.4.1 Condiciones de Calificación

Las 3 condiciones de calificación son: 1) Afinidad, 2) Slater, y 3) Mangasarian-Fromovitz. Para nuestro problema $\mathcal{PE}_T(\hat{x}_0)$ estas condiciones toman la forma:

1. **Afinidad.-** Todas las funciones F_k^i, $P_s^{j_s}$ y $Q_t^{i_t}$ son afines en sus dominios.
2. **Slater.-** La funciones F_k^i son afines, las funciones $P_s^{j_s}$ y $Q_t^{i_t}$ son convexas, y existe $\hat{w} \in W$ tal que:

   2.1 Para $i = 1, \cdots, n$ y $k = 0, \cdots, T-1$,
   $$F_k^i(\hat{w}) = 0.$$

   2.2 Para $s = 0, \cdots, T-1$, $t = 1, \cdots, T$, $j_s = 1, \cdots, n_s$ y $i_t = 1, \cdots, m_t$,
   $$P_s^{j_s}(\hat{w}) < 0, \qquad Q_t^{i_t}(\hat{w}) < 0.$$

3. **Mangasarian-Fromovitz.-** Dado $\hat{w} \in W$, satisfacerá la condición de calificación de Mangasarian-Fromovitz cuando:

   3.1 El conjunto de gradientes $\left\{\nabla_z F_k^i(\hat{w})\right\}_{\substack{i=1,\cdots,n,\\ k=0,\cdots,T-1.}}$ es **linealmente independiente**.

   3.2 Existe $d \in \mathbb{R}^{(n+m)T}$ tal que:

      a) Para $i = 1, \cdots, n$ y $k = 0, \cdots, T-1$,

$$\langle d, \nabla_w F_k^i(\hat{w})\rangle = 0.$$

      b) Para aquellos gradientes de restricciones activas

$$\langle d, \nabla_w P_s^{j_s}(\hat{w})\rangle < 0,$$

$$\langle d, \nabla_w Q_t^{i_t}(\hat{w})\rangle < 0,$$

      donde para $s = 0, \cdots, T-1$ y $t = 1, \cdots, T$,

$$\begin{array}{l} j_s \in J_s(\hat{w}) := \{j_s \in [1, \cdots, n_s] : P_s^{j_s}(\hat{w}) = 0\}, \\ i_t \in I_t(\hat{w}) := \{i_t \in [1, \cdots, m_t] : Q_t^{i_t}(\hat{w}) = 0\}. \end{array}$$

El problema que se nos presenta es cómo traducir estas condiciones en términos de las funciones f_t^0, f_t, p_t, q_t. En este sentido tenemos el siguiente resultado.

Lema 3.4.1 (Condiciones de calificación Afín y Slater) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ asumamos $(H1)_J - (H4)_J$ y (R). Entonces para el problema estático equivalente $\mathcal{PE}_T(\hat{x}_0)$ se cumplirá:*

1. *Si las funciones f_k^i, $p_s^{j_s}$ y $q_t^{i_t}$ son afines en sus dominios, entonces*

$$\mathcal{PE}_T(\hat{x}_0) \quad \textit{satisface la condición de afinidad.} \qquad (Afin)$$

2. *Si las funciones f_k^i son afines, las funciones $p_s^{j_s}$ y $q_t^{i_t}$ son convexas, y existe una sucesión de decisiones $(\tilde{u}_0, \cdots, \tilde{u}_{T-1})$, con su proceso asociado $(\tilde{x}_0, \cdots, \tilde{x}_T)$, tales que para $s = 1, \cdots, T$, $j_s = 1, \cdots, n_s$, $t = 0, \cdots, T-1$, y $i_t = 1, \cdots, m_t$,*

$$p_s^{j_s}(\tilde{x}_s) < 0 \quad y \quad q_t^{i_t}(\tilde{u}_t) < 0.$$

   Entonces,

$$\mathcal{PE}_T(\hat{x}_0) \quad \textit{satisface la condición de Slater.} \qquad (Slater)$$

Prueba. La demostración es inmediata, basta tener en cuenta las definiciones de las funciones restricciones del problema equivalente $\mathcal{PE}_T(\hat{x}_0)$, dados en el Lema 3.3.1.

$$\begin{aligned} F_t^i(w) &= x_{t+1}^i - f_t^i(x_t, u_t), \\ P_s^{j_s}(w) &= p_s^{j_s}(x_s), \\ Q_t^{i_t}(w) &= q_t^{i_t}(u_t). \end{aligned}$$

Se deduce claramente la afinidad o la convexidad de las funciones a partir de las hipótesis de la proposición para cada caso (Afinidad o Slater). ∎

Notemos que estas dos condiciones de calificación son globales es decir, si se satisfacen, entonces todo $w \in \operatorname{Adm}\mathcal{P}_{\mathrm{T}}(\hat{\mathrm{x}}_0)$ es un punto regular para KKT. De otro lado la condición de calificación de Mangasarian-Fromovitz es local. Es decir sólo nos permite ver si en un elemento $w \in \operatorname{Adm}\mathcal{P}_{\mathrm{T}}(\hat{\mathrm{x}}_0)$ podemos aplicar la condición de KKT.

Mangasarian-Fromovitz Para la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ con su proceso generado $(\hat{x}_0, \hat{x}_1, \cdots, \hat{x}_T)$, definamos $\Lambda = (\Lambda_0, \cdots, \Lambda_{T-1}) : (\mathbb{R}^m)^T \to (\mathbb{R}^n)^T$,

$$\Lambda(u_0, \cdots, u_{T-1}) := (x_1, \cdots, x_T),$$

donde $(x_1, \cdots, x_T)$ es la única solución del sistema $\Xi(u_0, \cdots, u_{T-1})$ (definido a partir de $(u_0, \cdots, u_{T-1})$)

$$\begin{aligned} x_1 &= D_u f_0(\hat{x}_0, \hat{u}_0) u_0, \\ x_{r+1} &= D_u f_r(\hat{x}_r, \hat{u}_r) u_r + D_x f_r(\hat{x}_r, \hat{u}_r) x_r, \qquad r = 2, \cdots, T. \end{aligned}$$

Con la ayuda de estas funciones tenemos el siguiente lema.

Lema 3.4.2 (Mangasarian-Fromovitz (MF)) *Para el problema* $\mathcal{P}_T(\hat{x}_0)$ *asumamos* $(H1)_J - (H4)_J$ *y* (R). *Consideremos una sucesión de decisiones* $(\hat{u}_0, \cdots, \hat{u}_{T-1}) \in \operatorname{Adm}\mathcal{P}_{\mathrm{T}}(\hat{\mathrm{x}}_0)$ *con su proceso asociado* $(\hat{x}_0, \hat{x}_1, \cdots, \hat{x}_T)$.

Si existen vectores $d_0, \cdots, d_{T-1} \in \mathbb{R}^m$, *tales que*

$$\begin{aligned} &\langle \nabla_{x_s} q_t^{i_t}(\hat{u}_t), d_t \rangle < 0, && \text{para } t = 1, \cdots, T, \\ & && \text{y } \ i_t \in I_t(\hat{u}_t), \\ &\langle \nabla_{x_s} p_s^{j_s}(\hat{x}_s), \Lambda_s(d) \rangle < 0, && \text{para } s = 0, \cdots, T-1, \\ & && \text{y } \ j_s \in J_s(\hat{x}_s), \end{aligned} \qquad (MF - PCD)$$

donde $d := (d_0, \cdots, d_{T-1}) \in (\mathbb{R}^m)^T$ *y para* $s = 0, \cdots, T-1$ *y* $t = 1, \cdots, T$,

$$\begin{aligned} J_s(\hat{x}_s) &= \{j_s \in [1, \cdots, n_s] \, : \, p_s^{j_s}(\hat{x}_s) = 0\}, \\ I_t(\hat{u}_t) &= \{i_t \in [1, \cdots, m_t] \, : \, q_t^{i_t}(\hat{u}_t) = 0\}. \end{aligned}$$

Entonces $\hat{w} := (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ satisface la condición de Mangasarian-Fromovitz para el problema estático equivalente $\mathcal{PE}_T(\hat{x}_0)$.

Prueba. Para demostrar este lema tenemos que analizar la condición de Mangasarian-Fromovitz en el problema equivalente estático $\mathcal{PE}_T(\hat{x}_0)$ para el punto $\hat{w} := (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$.

Para que este estudio sea lo más correcto posible reformulemos el problema $\mathcal{PE}_T(\hat{x}_0)$ de modo que la condición de Mangasarian-Fromovitz sea la que usualmente se cita en los libros.

En esta dirección usando las funciones del Lema 3.3.1 definamos para $t = 0, \cdots, T-1$, y $s = 1, \cdots, T$,

$$\begin{aligned}
F_t(w) := (F_t^1(w), \cdots, F_t^n(w)) &= x_{t+1} - f_t(x_t, u_t),\\
P_s(w) := (P_s^1(w), \cdots, P_s^{n_s}(w)) &= p_s(x_s),\\
Q_t(w) := (Q_t^1(w), \cdots, Q_t^{m_t}(w)) &= q_t(u_t).
\end{aligned}$$

A su vez con estas funciones definamos $F : W \to (\mathbb{R}^n)^T$, $P : W \to \prod_{s=1}^{T} \mathbb{R}^{n_s}$, $Q : W \to \prod_{t=0}^{T-1} \mathbb{R}^{m_t}$,

$$\begin{aligned}
F(w) &:= (F_0(w), \cdots, F_{T-1}(w)),\\
P(w) &:= (P_1(w), \cdots, P_T(w)),\\
Q(w) &:= (Q_0(w), \cdots, Q_{T-1}(w)).
\end{aligned}$$

Finalmente con P y Q definimos $G(w) := (P(w), Q(w))$. Con estas funciones el problema $\mathcal{PE}_T(\hat{x}_0)$ es equivalente a

$$\min_{\substack{F(w)=0,\\ G(w)\leq 0.}} -F^0(w). \tag{ϕ}$$

Restricciones Activas A partir de la definición de G, las restricciones activas para un punto $\hat{w} = (\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \hat{u}_{T-1})$ estarán determinadas por los conjuntos

$$\begin{aligned}
J_s(\hat{x}_s) &:= \{j_s \in [1, \cdots, n_s] \ : \ p_s^{j_s}(\hat{x}_s) = 0\},\\
I_t(\hat{u}_t) &:= \{i_t \in [1, \cdots, m_t] \ : \ q_t^{i_t}(\hat{u}_t) = 0\}.
\end{aligned}$$

donde $s = 0, \cdots, T-1$ y $t = 1, \cdots, T$.

Para el problema $\mathcal{PE}_T(\hat{x}_0)$ reformulado con estas funciones (ϕ), la condición de Mangasarian-Fromovitz en $w = (x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$ es:
"Existe $d \in \ker DF(\hat{w})$ tal que

$$\langle d, \nabla_w G_j(\hat{w}) \rangle < 0, \tag{$(MF)_E$}$$

para las **restricciones activas***".*

Por lo tanto nuestro lema será demostrado si encontramos la forma que tiene $\ker DF(\hat{w})$, en relación a las funciones f_t y f_t^0. Y finalmente si hacemos un estudio similar de los gradientes de las coordenadas de la función G, para las restricciones activas.

$\ker \mathbf{DF}(\hat{\mathbf{w}})$ De la definición de la función F no es difícil calcular su matriz derivada $DF(w) \in \mathbb{R}^{nT\times(m+n)T}$,

$$\left(\begin{array}{ccccc|cccc}
I & 0 & \cdots & 0 & 0 & -D_u f_0 & 0 & \cdots & 0 \\
-D_x f_1 & I & \dots & 0 & 0 & 0 & -D_u f_1 & \cdots & 0 \\
0 & -D_x f_2 & \dots & 0 & 0 & 0 & 0 & \ddots & \vdots \\
0 & 0 & \ddots & \vdots & \vdots & 0 & 0 & \ddots & \vdots \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \dots & I & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & \dots & -D_x f_{T-1} & I & 0 & 0 & \cdots & -D_u f_{T-1}
\end{array}\right)$$

Por lo que para $\hat{w} := (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$, un vector

$$(x_1, x_2, \cdots, x_T, u_0, u_1, \cdots, u_{T-1}) \in KerDF(\hat{w}),$$

si y sólo si es solución del sistema Ξ (asociado a $\hat{w}$):

$$\left|\begin{array}{rcl}
x_1 & = & D_u f_0(\hat{x}_0, \hat{u}_0)u_0 \\
x_2 & = & D_u f_1(\hat{x}_1, \hat{u}_1)u_1 + D_x f_1(\hat{x}_1, \hat{u}_1)x_1 \\
 & & \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\
x_T & = & D_u f_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1})u_{T-1} + D_x f_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1})x_{T-1}.
\end{array}\right.$$

Del cual se deduce que para cada $(u_0, u_1, \cdots, u_{T-1})$, existe un único proceso $(x_1, u_1, \cdots, x_T)$, que es solución del sistema $\Xi(u_0, \cdots, u_{T-1})$,

$$\begin{array}{rcll}
x_1 & = & D_u f_0(\hat{x}_0, \hat{u}_0)u_0, & \\
x_{r+1} & = & D_u f_r(\hat{x}_r, \hat{u}_r)u_r + D_x f_r(\hat{x}_r, \hat{u}_r)x_r, & \qquad r = 2, \cdots, T.
\end{array}$$

Por lo que podemos construir la función $\Lambda = (\Lambda_0, \cdots, \Lambda_{T-1}) : (\mathbb{R}^m)^T \to (\mathbb{R}^n)^T$, definida anteriormente.

De otro lado usando el Teorema del Núcleo e Imagen del Algebra Lineal (ver [16, p.74]) y el Lema 3.2.1 (independencia lineal de las gradientes) deducímos que el rango según filas de la matriz es nT, por ello también lo es según sus columnas, consecuentemente la dimensión de la Imagen es nT. Tenemos

$$\begin{array}{rcl}
\dim ImDF(z)^T & = & nT, \\
\dim(\ker DF(z)) + \dim ImDF(z)^T & = & (n+m)T.
\end{array}$$

Entonces $\dim(\ker DF(z)) = mT$. Con estos resultados podemos caracterizar el Núcleo de $DF(\hat{w})$, según

$$v = (v_0, v_1) \in Ker DF(\hat{w}) \subseteq \mathbb{R}^{mT} \times \mathbb{R}^{nT} \text{ si y sólo si } v_1 \text{ es solución de } \Xi(v_0).$$

Además es claro que una base del núcleo se obtendrá tomando los elementos de la base canónica de $e_i \in \mathbb{R}^{mT}$ para generar vectores $(e_i, \Lambda(e_i))$.

Equivalente de la Condición M-F De la caracterización de $\ker DF(\hat{w})$, la condición de Mangasarian-Fromovitz para $\hat{w}$, $((MF)_E)$, será equivalente a (MF-PCD) del Lema 3.4.2. ∎

Una vez estudiado como interpretar las condiciones de calificación para el problema de control $\mathcal{P}_T(\hat{x}_0)$, tenemos el siguiente teorema.

Teorema 3.4.1 (Principio Débil de Pontryagin-KKT (PdPKKT)) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ asumanos $(H1)_J - (H4)_J$ y (R). Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso asociado $(\hat{x}_0, \cdots, \hat{x}_T)$, una* **solución óptima**.

Si el problema satisface cualquiera de las condiciones dadas en el Lema 3.4.1 (Afinidad o Slater) o en el Lema 3.4.2 (Mangasarian-Fromovitz) para $(\hat{u}_0, \cdots, \hat{u}_{T-1})$.

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, $\{\hat{\nu}_s \in \mathbb{R}^{n_s}_+ : s = 1, \cdots, T\}$, $\{\hat{\mu}_t \in \mathbb{R}^{m_t}_+ : t = 0, \cdots, T-1\}$ **no todos nulos**, *tales que:*

1. *Para $t = 1, \cdots, T-1$,*

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t, \hat{\nu}_t), \qquad (EA)_K$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T}\left(f^0_T(\hat{x}_T) - \langle \hat{\nu}_T, p_T(\hat{x}_T)\rangle\right).$

2. *Para $t = 1, \cdots T$ y $j_t = 1, \cdots, n_t$,*

$$\hat{\nu}^{j_t}_t p^{j_t}_t(\hat{x}_t) = 0. \qquad (C(x))_K$$

(Condición de complementariedad para x_t).

3. *Para $t = 0, \cdots, T-1$,*

$$\nabla_{u_t}\Big(H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}(t), \hat{\nu}_t) - \langle \hat{\mu}_t, q_t(\hat{u}_t)\rangle\Big) = 0. \qquad (PdPK)$$

4. *Para $t = 0, \cdots T-1$ y $i_t = 1, \cdots, m_t$,*

$$\hat{\mu}^{i_t}_t q^{i_t}_t(\hat{u}_t) = 0. \qquad (C(u))_K$$

(Condición de complementariedad para u_t).

Por lo tanto, de los items 3 y 4, para $t = 0, \cdots, T-1$, $\hat{u}_t$ es un **punto crítico** *del problema*

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, a_0, \hat{\psi}_t, \hat{\nu}_t), \qquad (PMdPK)_t$$

y de los items 1 y 2, los vectores $\hat{\psi}_t$ son las variables adjuntas.

Donde $H_t(x, u, \psi, \nu) = f_t^0(x, u) + \langle \psi, f_t(x, u) \rangle - \langle \nu, p_t(x) \rangle$ son los hamiltonianos.

Prueba. Del Lema 3.3.1 sabemos que bajo las hipótesis $(H1)_J - (H4)_J$ y (R) $\mathcal{P}_T(\hat{x}_0)$ es equivalente al problema estático $\mathcal{PE}_T(\hat{x}_0)$.

$$\max_{\substack{F_k^i(w)=0,\\ P_s^{js}(w)\leq 0,\\ Q_t^{it}(w)\leq 0.}} F^0(w) \quad \equiv \quad \min_{\substack{F_k^i(w)=0,\\ P_s^{js}(w)\leq 0,\\ Q_t^{it}(w)\leq 0.}} -F^0(w).$$

La demostración de este teorema es consecuencia del Teorema 3.3.1 (Principio Débil de Pontryagin-John, PdPJ) pues en vista de los Lemas 3.4.1 y 3.4.2, $\hat{w} := (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ satisface una de las condiciones de calificación (Afinidad, Slater o MF) para $\mathcal{PE}_T(\hat{x}_0)$, lo que nos asegura que $a_0 > 0$, por lo cual normalizando los multiplicadores tenemos $a_0 = 1$, y consecuentemente deducímos el presente teorema. ■

Estudiemos con detenimiento algunas consecuencias del resultado anterior. Primero un resultado análogo al Corolario 3.3.1 para "estados interiores" del Teorema 3.3.1 (PdPJ).

Corolario 3.4.1 (Estados interiores) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ asumanos $(H1)_J - (H4)_J$ y (R). Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso asociado $(\hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima**.

Si el problema satisface cualquiera de las condiciones dadas en el Lema 3.4.1 (Afinidad o Slater) o en el Lema 3.4.2 (MF) para $(\hat{u}_0, \cdots, \hat{u}_{T-1})$. Supongamos además que $\hat{x}_t \in \operatorname{int} X_t$, para $t = 1, \cdots, T$.

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, $\{\hat{\mu}_t \in \mathbb{R}_+^{m_t} : t = 0, \cdots, T-1\}$ **no todos nulos**, *tales que:*

1. *Para $t = 1, \cdots, T-1$,*

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \qquad (EA)_K$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T} f_T^0(\hat{x}_T)$.

2. *Para $t = 0, \cdots, T-1$,*

$$\nabla_{u_t} \Big(H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) - \langle \hat{\mu}_t, q_t(\hat{u}_t) \rangle \Big) = 0. \qquad (PdPK)$$

3. *Para* $t = 0, \cdots T-1$ *y* $i_t = 1, \cdots, m_t$,

$$\hat{\mu}_t^{i_t} q_t^{i_t}(\hat{u}_t) = 0. \qquad (C(u))_K$$

(Condición de complementariedad para u_t*).*

Por lo tanto para $t = 0, \cdots, N-1$, $\hat{u}_t$ *es punto crítico del problema*

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t),$$

donde $H_t(x, u, \psi) := f_t^0(x, u) + \langle \psi, f_t(x, u) \rangle$ *es el hamiltoniano.*

Prueba. En el Teorema 3.4.1 (PdPKKT), desde la interioridad de $x_t \in \operatorname{int} X_t$, se deduce que no tiene restricciones activas respecto a "x". Por lo tanto desde la complementariedad $(C(x))_K$, se deduce que para todo multiplicador $\nu_t^{j_t} = 0$ en el Teorema 3.4.1. De ello el hamiltoniano se reduce a

$$H_t(x_t, u_t, \psi_t, \nu_t) = H_t(x_t, u_t, \psi_t) = f_t^0(x, u) + \langle \psi_t, f_t(x, u) \rangle.$$

Todas las otras consecuencias de este corolario salen simplemente reemplazando el hamiltoniano por éste más simple, teniendo en cuenta que $\nu_t^{j_t} = 0$. ■

En el Teorema 3.4.1 vemos que para cada $t = 0, \cdots, T-1$, $\hat{u}_t$ es un punto crítico KKT para el subproblema

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t, \hat{\nu}_t). \qquad (PMdPK)_t$$

De otra parte sabemos que bajo ciertas condiciones especiales (Convexidad) el hecho de ser punto crítico nos asegura que es un óptimo (condición suficiente). En esta dirección tenemos el siguiente corolario.

Corolario 3.4.2 (PFP - Convexidad) *Para el problema* $\mathcal{P}_T(\hat{x}_0)$ *asumamos* $(H1)_J - (H4)_J$, (R), *y además las siguientes hipótesis:*

- *Para* $t = 0, \cdots, T-1$, *sea* $x_t \in X_t$, *fijo y arbitrario, entonces*

$$f_t(x_t, \cdot) : B_t \to \mathbb{R}^n \qquad \text{es afín.} \qquad (Afin - u)$$

- *Para* $t = 0, \cdots, T-1$, *sea* $x_t \in X_t$, *fijo y arbitrario, entonces*

$$f_t^0(x_t, \cdot) : B_t \to \mathbb{R}^n \qquad \text{es cóncava.} \qquad (Conc - u)$$

- *Para* $t = 0, \cdots, T-1$, B_t *convexa y*

$$q_t^{i_t} : B_t \to \mathbb{R} \qquad \text{es convexa.} \qquad (Cvxa - u)$$

Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, *con su proceso* $(\hat{x}_1, \cdots, \hat{x}_T)$, *una* **solución óptima**.

Si el problema satisface cualquiera de las condiciones dadas en el lema 3.4.1 (calificación afín o de Slater) o en el lema 3.4.2 (Mangasarian-Fromovitz) para $(\hat{u}_0, \cdots, \hat{u}_{T-1})$.

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, $\{\hat{\nu_s} \in \mathbb{R}_+^{n_s} : s = 1, \cdots, T\}$ **no todos nulos**, *tales que:*

1. *Para* $t = 1, \cdots, T-1$,

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t; \hat{\psi}_t, \hat{\nu}_t), \qquad (EA)_K$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T} \left(f_T^0(\hat{x}_T) - \langle \hat{\nu}_T, p_T(\hat{x}_T) \rangle \right)$.

2. *Para* $t = 1, \cdots T$ *y* $j_t = 1, \cdots, n_t$,

$$\hat{\nu}_t^{j_t} p_t^{j_t}(\hat{x}_t) = 0. \qquad (C(x))_K$$

(condición de complementariedad para x_t*).*

3. *Para* $t = 0, \cdots, T-1$, $\hat{u}_t$ *es* **solución óptima** *del problema*

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t, \hat{\nu}_t), \qquad (PMdPK)_t$$

donde $H_t(x, u, \psi, \nu) = f_t^0(x, u) + \langle \psi, f_t(x, u) \rangle - \langle \nu, p_t(x) \rangle$ *son los hamiltonianos.*

Prueba. La demostración se sigue del Teorema 3.4.1, notando que bajo las hipótesis (Afin, Conc-u,Cvxa-u) los problemas $(PMdPK)_t$

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t, \hat{\nu}_t), \qquad (PMdPK)_t$$

son programas convexos. Por lo tanto la condición de KKT es suficiente. Por su parte del Teorema 3.4.1, $\hat{u}_t$ es punto crítico de este problema luego desde la condición suficiente de KKT , es una **solución óptima.** ■

Juntando ambos corolarios tenemos un primer resultado que coincide plenamente con el Principio Fuerte de Pontryagin (Proposición 3.1.2).

Corolario 3.4.3 (PFP - Convexidad e Interioridad) *Para* $\mathcal{P}_T(\hat{x}_0)$ *asumamos las hipótesis* $(H1)_J - (H4)_J$, (R), *y también las del Corolario 3.4.2.*

Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, *con su proceso* $(\hat{x}_1, \cdots, \hat{x}_T)$, *una* **solución óptima** *que satisface además para* $t = 1, \cdots, T$,

$$\hat{x}_t \in \operatorname{int} X_t. \qquad (I)$$

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, **no todos nulos**, *tales que:*

1. Para $t = 1, \cdots, N-1$,

$$\psi_{t-1} = \nabla_x H(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \tag{EA}$$

con condición final $\psi_{N-1} = \nabla_x f_T^0(\hat{x}_T)$.

2. Para cada $t = 0, \cdots, N-1$, $\hat{u}_t$ es solución óptima del problema

$$\max_{u_t \in U_t} H(\hat{x}_t, u_t, \hat{\psi}_t). \tag{$(PM)_t$}$$

Donde las funciones $H_t(x, u, \psi) = f_t^0(x, u) + \langle \psi, f_t(x, u) \rangle$ son los hamiltonianos.

Prueba. La prueba es inmediata usando los dos corolarios anteriores. ■

Observación. Este es el primer resultado en el que vemos que es cierto el Principio Fuerte del Máximo de Pontryagin. En el siguiente capítulo estudiaremos otros casos en los cuales es cierto, basándonos en otras consideraciones teóricas.

Finalmente explotando al máximo las condiciones de KKT usemos la condición suficiente de optimalidad aplicada al problema equivalente estático $\mathcal{PE}_T(\hat{x}_0)$.

Corolario 3.4.4 (PFP - Suficiente) *Para $\mathcal{P}_T(\hat{x}_0)$ asumamos $(H1)_J$–$(H4)_J$, (R), y además:*

- *Para $t = 0, \cdots, T-1$,*

$$f_t : A_t \times B_t \to \mathbb{R}^n \qquad \text{es afín.} \tag{$Afin$}$$

- *Para $t = 0, \cdots, T-1$, sea x_t, fijo y arbitrario, entonces*

$$f_t^0(x_t, \cdot) : B_t \to \mathbb{R}^n \qquad \text{es cóncava.} \tag{$Conc-u$}$$

- *Para $t = 0, \cdots, T-1$, y $j_s = 1, \cdots, n_s$, sea A_t convexa y*

$$p_s^{j_s} : A_s \to \mathbb{R} \qquad \text{es convexa.} \tag{$Cvxa-x$}$$

- *Para $t = 0, \cdots, T-1$, y $i_t = 1, \cdots, m_t$, sea B_t convexa y*

$$q_t^{i_t} : B_t \to \mathbb{R} \qquad \text{es convexa.} \tag{$Cvxa-u$}$$

Supongamos que el problema satisface cualquiera de las condiciones dadas en el Lema 3.4.1 (Afinidad o Slater) o en el Lema 3.4.2 (MF) para $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso $(\hat{x}_1, \cdots, \hat{x}_T)$. Y que, asociados a la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ existan vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, $\{\hat{\nu}_s \in \mathbb{R}_+^{n_s} : s = 1, \cdots, T\}$ **no todos nulos**, *tales que:*

1. *Para* $t = 1, \cdots, T-1$,

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t, \hat{\nu}_t), \qquad (EA)_K$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T} \left(f_T^0(\hat{x}_T) - \langle \hat{\nu}_T, p_T(\hat{x}_T) \rangle \right)$.

2. *Para* $t = 1, \cdots T$ *y* $j_t = 1, \cdots, n_t$,

$$\hat{\nu}_t^{j_t} p_t^{j_t}(\hat{x}_t) = 0. \qquad (C(x))_K$$

(Condición de complementariedad para x_t*).*

3. *Para* $t = 0, \cdots, T-1$, $\hat{u}_t$ *es* **solución óptima** *del problema*

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t, \hat{\nu}_t). \qquad (PMdPK)_t$$

Entonces la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$*, con su proceso* $(\hat{x}_1, \cdots, \hat{x}_T)$*, es una* **solución óptima global** *del problema* $\mathcal{P}_T(\hat{x}_0)$*.*

Prueba. La demostración se obtiene simplemente observando que con las hipótesis hechas (Afín, Concva-u, Cvxa-x, Cvxa-u) el problema equivalente

$$\min_{\substack{F_k^i(w)=0,\\ P_s^{j_s}(w)\leq 0,\\ Q_t^{i_t}(w)\leq 0.}} -F^0(w)$$

es un programa convexo.

Sabemos que al satisfacer $\hat{w} = (\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ la condición de calificación se deduce que $\hat{w} = (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ es un punto crítico (punto KKT) de éste problema. Por lo cual de la condición suficiente KKT se deduce que $\hat{w}$ es un óptimo global del problema $\mathcal{PE}_T(\hat{x}_0)$. Por lo cual la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ es una solución óptima para el problema $\mathcal{P}_T(\hat{x}_0)$. ■

En la siguiente sección buscaremos extender aún más el Principio Débil, para el caso en el que ya no se conoce la forma de U_t (es decir ya no se la puede definir usando una función q_t). Para dar la condición de optimalidad en este caso lo que haremos será estudiar el cono Bouligand y la condición necesaria primal, desde la perspectiva de la teoría de optimalidad desarrollada por V.G. Boltayanskii.

3.5 Principio Débil: Boltyanskii

Aún cuando el resultado de la sección anterior es bastante general, exige una estructura particular para los conjuntos X_s, U_t. (es decir, que estan determinados por una función p_s, q_t respectivamente) ¿Cómo podríamos estudiar el caso en el que sólo se conoce la "forma geométrica" de U_t?

Sabemos que el Teorema de Condición Necesaria Primal (ver [15, p.44]) no requiere conocer exactamente la forma de U_t sino que basta conocer el cono Bouligand (ver [15, p.25]). Según este teorema $\hat{u}_t$ será un punto crítico para el problema

$$\max_{u \in U_t} H_t(\hat{x}_t, u, \hat{\psi}_t),$$

cuando para todo $d \in B_{u_t}(\hat{u}_t)$,

$$\langle \nabla_u H_t(\hat{x}_t, \hat{u}_t, \hat{\lambda}_t), d \rangle \leq 0.$$

Pues bien, usando la Teoría de Boltyanskii se puede obtener una primera aproximación en este sentido, demostrándose para un contexto más general el Principio Débil de Pontryagin.

Para realizar esta aproximación correctamente necesitamos que nuestro problema $\mathcal{P}_T(\hat{x}_0)$ satisfaga las siguientes hipótesis.

- Para $t = 0, \cdots, T$, y $i = 0, \cdots, n$,

$$f_t^i \qquad \text{son funciones de tipo } C^1 \text{ en sus respectivos dominios.} \qquad (H1)_B$$

- Para $t = 1, \cdots, T$,

$$x_t \in \operatorname{int} X_t. \qquad (I)_B$$

Los conceptos y los teoremas que se usan en esta sección, forman parte de la Teoría de Boltyanskii, y se puede encuentrar un resúmen en el Apéndice B. Uno de estos conceptos es el de cúpula de un conjunto en un punto. Este concepto es menos general que el de cono Bouligand, pero en muchos casos es equivalentes, por lo cual el teorema que aquí demostraremos es bastante general.

Del mismo modo que procedimos en los teoremas anteriores (Teoremas 3.2.1, 3.3.1) primeramente transformaremos nuestro problema dinámico en un problema estático.

Lema 3.5.1 (Transformación: dinámico a estático) *El problema de control de procesos discretos en horizonte finito $\mathcal{P}_T(\hat{x}_0)$ es equivalente a un problema de optimización estático $\mathcal{PE}_{\mathcal{B}T}(\hat{x}_0)$.*

Prueba. Denotemos $W = \prod_{s=1}^{T} A_s \times \prod_{t=0}^{T-1} B_t \subset \mathbb{R}^{(n+m)T}$ y a sus elementos como $w := (x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$.

Definamos $F^0 : W \to \mathbb{R}$, por

$$F^0(z) = \sum_{k=0}^{T-1} f_k^0(x_k, u_k) + f_T^0(x_T),$$

y

$$\Omega = \left\{ w \in W \;\middle|\; \begin{array}{l} \text{para } \; i = 1, \cdots, n, \quad \text{y} \\ t = 0, \cdots, T-1, \; F_t^i(w) = 0 \end{array} \right\},$$

donde, para $t = 0, \cdots, T-1$, y $i = 1, \cdots, n$,

$$F_t^i(w) = x_{t+1}^i - f_t^i(x_t, u_t).$$

Sea además,

$$\Xi := \left\{ w = (x_1, \cdots, x_T, u_0, \cdots, u_{T-1}) \in W \;\middle|\; \begin{array}{c} u_t \in U_t, \; x_{t+1} \in X_{t+1}, \\ \text{para } t = 0, \cdots, T-1. \end{array} \right\}.$$

Con estas definiciones es claro que el problema de control discreto $\mathcal{P}_T(\hat{x}_0)$ es equivalente al problema de optimización estática con restricciones,

$$\mathcal{P}_T(\hat{x}_0) \quad \equiv \quad \min_{\Omega \cap \Xi} -F^0. \qquad \mathcal{PE}_{\mathcal{B}T}(\hat{x}_0)$$

■

Teorema 3.5.1 (Principio Débil de Pontryagin-Boltyanskii (PdPB)) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ asumamos $(H1)_B$ y $(I)_B$. Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso $(\hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima**.

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, **no todos nulos** *tales que:*

1. *Para $t = 1, \cdots, N-1$,*

$$\hat{\psi}_{t-1} = \nabla_x H(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \qquad (EA)$$

con condición final, $\hat{\psi}_{N-1} = \nabla_x f_T^0(\hat{x}_T)$.

2. *Para $t = 0, \cdots, N-1$, se cumple: Para cualquier $\delta u \in L_t$,*

$$\langle \delta u \, ; \; \nabla_u H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) \rangle \leq 0, \qquad (PdPB)_t$$

donde para $t = 0, \cdots, T-1$, L_t es la cúpula de $\hat{u}_t$ respecto al conjunto U_t.

Desde **2**. *deducímos que* $\hat{u}_t$ *es punto crítico del problema*

$$\max_{u \in U_t} H(\hat{x}_t, u, \hat{\psi}_t), \qquad (PM)_t$$

donde las funciones $H_t(x_t, u_t, \psi_t) = f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$ *son los hamiltonianos.*

Prueba. En primer lugar calculemos para $\hat{w} = (\hat{x}_1, \cdots, \hat{x}_{T-1}, \hat{u}_0, \cdots, \hat{u}_{T-1})$ su cúpula Ξ para luego usar la condición necesaria de optimalidad de Boltyanskii.

Cúpula de Ξ Definamos el conjunto

$$L = \left\{ (x_1, \cdots, x_T, u_0, \cdots, u_{T-1}) \in \mathbb{R}^{nT} \times \mathbb{R}^{mT} : u_t \in L_t \text{ para } t = 0, \cdots, T-1 \right\}.$$

Demostremos que L es cúpula de Ξ en $\hat{w}$, y además

$$\delta z = (\delta x_1, \cdots, \delta x_T, \delta u_0, \cdots, \delta u_{T-1}) \in L \Leftrightarrow \delta u_t \in L_t \text{ para } t = 0, \cdots, T-1.$$

Para probar esto procedamos según la definición de la cúpula (ver Apéndice B)

1. L es un cono convexo, pues es claro que desde su definición

$$L = (\mathbb{R}^n)^T \times L_1 \times \cdots \times L_T.$$

2. Probemos la siguiente condición de cúpula. Desde que L_t es cúpula de $\hat{u}_t$ en U_t.

Existen entornos $B_t(0; \epsilon_t)$ y funciones continuas $\phi_t : L_t \cap B_t(0; \epsilon_t) \to \mathbb{R}^m$ satisfaciendo:

(1) $\phi_t(L_t \cap B_t(0; \epsilon_t)) \subset U_t$ y $\phi_t(0) = \hat{u}_t$.

(2)

$$\lim_{\substack{h \to 0 \\ h \in L_t}} \frac{\|\phi_t(h) - \hat{w}\|}{|h|} = 0.$$

Tomemos $\epsilon > 0$ de modo que $B(\hat{w}; \epsilon) \subset (\mathbb{R}^n)^T \times (\mathbb{R}^m)^T$, satisfaga

$$\pi_{Y_t}(B(\hat{w}; \epsilon)) \subseteq B_t(0; \epsilon_t) \subset \mathbb{R}^m,$$

donde π_{Y_t} es la proyección, considerando los elementos de $(\mathbb{R}^n)^T \times (\mathbb{R}^m)^T$, como $(x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$, definamos la función

$$\begin{array}{llll} \phi : & B(\hat{z}; \epsilon) \cap L & \to & \Xi \\ & (x_1, \cdots, x_T, u_0, \cdots, u_{T-1}) & \mapsto & (x_1, \cdots, x_T, \phi_1(u_0), \cdots, \phi_{T-1}(u_{T-1})). \end{array}$$

Esta función cumple las condiciones necesarias para la definición de cúpula, desde que las cumplen las funciones ϕ_t. Ahora pasemos a aplicar la condición necesaria de Boltyanskii.

Condición Necesaria de Optimalidad de Boltyanskii Aplicando la Condición Necesaria de Optimalidad de Boltyanskii (Teorema B.2.4) al problema equivalente $\mathcal{PE}_{\mathcal{B}T}(\hat{x}_0)$ para $\hat{w}$.

Existen números $\mu_0 \leq 0$, $\{\bar{\psi}_t^i \in \mathbb{R} :$ con $1, \cdots, n$ y $t = 0, \cdots, T-1\}$ **no todos nulos** tales que:

Para todo $(\delta x_1, \cdots, \delta x_T, \delta u_1, \cdots, \delta u_T) \in L$,

$$\sum_{s=1}^{T} \left\langle \left(-\mu_0 \nabla_{x_s} F^0(\hat{w}) + \sum_{\alpha=1}^{n} \sum_{\gamma=0}^{T-1} \bar{\psi}_\alpha^\gamma \nabla_{x_s} F_\gamma^\alpha(\hat{w}) \right); \delta x_s \right\rangle + \\ + \sum_{t=0}^{T-1} \left\langle \left(-\mu_0 \nabla_{u_t} F^0(\hat{w}) + \sum_{\alpha=1}^{n} \sum_{\gamma=0}^{T-1} \bar{\psi}_\alpha^\gamma \nabla_{u_t} F_\gamma^\alpha(\hat{w}) \right); \delta u_t \right\rangle \leq 0. \quad (\beta)$$

A partir de estas dos enunciados se deduce nuestro teorema.

Sistema Adjunto Desde que (β) se cumple para todo vector de L. Consideremos algunos vectores particulares. Tenemos por ejemplo:

Para todo $\delta x_t \neq 0 \in \mathbb{R}^n$, con $t = 1, \cdots, T$,

$$(\underline{0}_1, \cdots, \underline{0}_{t-1}, \delta x_t, \underline{0}_{t+1}, \cdots, \underline{0}_T; \underline{0}_1, \cdots, \underline{0}_T) \in L.$$

Por lo cual desde (β) obtenemos que para $t = 1, \cdots, T$,

$$-\mu_0 \nabla_{x_t} F^0(\hat{w}) + \sum_{\alpha=1}^{n} \sum_{\gamma=0}^{T-1} \bar{\psi}_\alpha^\gamma \nabla_{x_t} F_\gamma^\alpha(\hat{w}) = 0 \in \mathbb{R}^n.$$

Si denotamos $\psi_\gamma = (\psi_1^\gamma, \ldots, \psi_n^\gamma) \in \mathbb{R}^n$, podemos simplificar la igualdad anterior como

$$-\mu_0 \nabla_{x_t} F^0(\hat{w}) + \sum_{\gamma=0}^{T-1} \psi_\gamma \circ D_{x_t} F_\gamma(\hat{w}) = 0. \quad (1)$$

Cálculo de Gradientes De los cálculos, para $t = 0, \cdots, T-1$,

$$\begin{array}{ll} D_{x_t} F_\gamma(\hat{w}) = \ 0, & \text{si } \gamma \notin \{t-1, t\}, \\ D_{x_t} F_\gamma(\hat{w}) = -I, & \text{si } \gamma = t-1, \\ D_{x_t} F_\gamma(\hat{w}) = D_{x_t} f_t(\hat{x}_t, \hat{u}_t), & \text{si } \gamma = t. \end{array}$$

Para $t = T$ se tiene

$$\begin{aligned} -\nabla_{x_T} F^0(\bar{w}) &= -\nabla_{x_T} f_T^0(\hat{x}_T) \in \mathbb{R}^n, \\ \nabla_{x_T} F_{T-1}(\hat{w}) &= -I \ \in \mathbb{R}^{n \times n}. \end{aligned}$$

Con estos cálculos en (1) obtenemos para $t = 1, \cdots, T-1$,

$$-\mu_0 \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) - \hat{\psi}_{t-1} \circ (I) + \hat{\psi}_t \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t) = 0, \quad (2)$$

lo cual es equivalente a

$$-\mu_0 \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) + \psi_t \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t) = \psi_{t-1}. \qquad (\phi)$$

De otro lado para $t = T$ en (1) tendríamos:

$$-\mu_0 \nabla_{x_T} f_T^0(\hat{x}_T) - \hat{\psi}_{T-1} \circ (I) = 0.$$

Por lo tanto: $\hat{\psi}_{T-1} = -\mu_0 \nabla_{x_T} f_T^0(\hat{x}_T)$.

Sabemos que $\mu_0 \leq 0$, demostremos que $\mu_0 < 0$. Supongamos por contradicción que $\mu_0 = 0$, entonces $\hat{\psi}_{T-1} = \mu_0 \nabla_{x_T} f_T^0(\hat{x}_T) = 0$. Además, desde (ϕ), tenemos

$$-\mu_0 \nabla_{x_{T-1}} f_{T-1}^0(\hat{x}_{T-1}, \hat{u}_{T-1}) - \psi_{T-1} \circ D_{x_{T-1}} f_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1}) = \hat{\psi}_{T-2},$$

por ello $\psi_{T-2} = 0$. Similarmente con $\mu_0 = 0$ y $\psi_{T-2} = 0$ se deduce que $\psi_{T-3} = 0$, así sucesivamente tendríamos $\mu_0 = \hat{\psi}_0 = \cdots = \hat{\psi}_{T-1} = 0$, lo que contradeciría el hecho de que son **no nulos**. Por ello $\mu_0 < 0$.

Para $t = 1, \cdots, T-1$, redefinamos $\hat{\psi}_t := -\mu_0^{-1}.\hat{\psi}_t$. Reemplazando en (ϕ), tenemos el sistema:

Para $t = 0, \cdots, T-1$,

$$\hat{\psi}_{t-1} = \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) + \hat{\psi}_t \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t), \qquad (\phi)$$

con condición final $\hat{\psi}_{T-1} = \nabla_{x_T} f_T^0(\hat{x}_T)$.

Usando el hamiltoniano $H_t(x_t, u_t, \psi_t) = f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$ tenemos

$$\nabla_x H_t(x_t, u_t, \psi_t) = \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) + \hat{\psi}_t \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t).$$

En consecuencia obtenemos, desde (ϕ), el Sistema Adjunto (EA) de nuestro teorema.

Principio Débil Del cálculo, tenemos para $t = 0, \cdots, T-1$,

$$\begin{aligned} -\nabla_{u_t} F^0(w) &= -\nabla_{u_t} f_t^0(x_t, u_t), \\ D_{u_t} F_t(w) &= D_{u_t} f_t(x_t, u_t). \end{aligned}$$

En (β), usemos los vectores $(\underline{0}_1, \cdots, \underline{0}_T; \underline{0}_1, \cdots, \underline{0}_{t-1}, \delta u_t, \underline{0}_{t+1}, \cdots, \underline{0}_T) \in L$, donde $\delta u_t \in L_t$, obtenemos

$$\left\langle \left(-\mu_0 \nabla_{u_t} f_t^0(x_t, u_t) + \hat{\psi}_t \circ D_{u_t} f_t(x_t, u_t) \right); \delta u_t \right\rangle \leq 0.$$

Considerando la redefinición de $\hat{\psi}_t$ deducímos para $t = 0, \cdots, T-1$,

$$\left\langle \nabla_{u_t} f_t^0(x_t, u_t) + \hat{\psi}_t \circ D_{u_t} f_t(x_t, u_t) \; ; \; \delta u_t \right\rangle \leq 0.$$

De lo cual se obtiene $(PdPB)_t$. ■

Observación. Este teorema logra generalizar el Teorema 3.4.1 (PdPKKTD). Pues ahora se usa la Condición Necesaria Primal (ver [15, p.44]). Sin embargo la generalización no es del todo completa, pues la hipótesis $(H1)_B$ es más restrictiva que las hipótesis $(H1)_J$ y $(H2)_J$.

Procediendo de un modo enteramente similar al teorema anterior podemos encontrar también una condición necesaria para el problema de control discreto con extremo derecho fijo (con un objetivo final). Es decir cuando $X_T = \{\hat{x}_T\}$ (ver Capítulo 1)

Corolario 3.5.1 (PdPB - Extremo Fijo) *Para el problema $\mathcal{P}_T(\hat{x}_0; \hat{x}_T)$ asumamos $(H1)_B$. Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso $(\hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima** *que satisface $(I)_B$.*

Entonces existen $a_0 \in \mathbb{R}$ y vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$ tales que:

1. $a_0 \geq 0$.

2. $(a_0, \hat{\psi}_0, \cdots, \hat{\psi}_{T-1})$ *es* **no nulo**.

3. *Para $t = 1, \cdots, T-1$,*

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t). \tag{EA}$$

4. *Para todo $\delta w_t \in L_t$:*

$$\left\langle \delta w_t \ ; \ \nabla_u H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t) \right\rangle \leq 0, \tag{$(PdPB)_t$}$$

donde $t = 0, \cdots, T-1$, L_t es la cúpula de $\hat{u}_t$, y $H_t(x_t, u_t, a_0, \psi_t) = a_0 f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$ es el Hamiltoniano.

Prueba. Vamos a proceder análogamente al resultado anterior es decir transformaremos el problema dinámico en un problema estático, notando que en este caso tenemos una variable menos ($x_T = \hat{x}_T$, constante), en consecuencia $W := A_1 \times \cdots \times A_{T-1} \times B_0 \times \cdots \times B_{T-1}$, (sus elementos serán $w = (x_1, \cdots, x_{T-1}, u_0, \cdots, u_{T-1})$).

Definamos $F^0 : W \to \mathbb{R}$, por

$$F^0(z) = \sum_{k=0}^{T-1} f_k^0(x_k, u_k) + f_T^0(\hat{x}_T),$$

y

$$\Omega = \left\{ w \in W \;\middle|\; \begin{array}{l} \text{para } \ i = 1, \cdots, n, \quad \text{y} \\ t = 0, \cdots, T-1, \ F_t^i(w) = 0 \end{array} \right\},$$

donde, para $t = 0, \cdots, T-2$, y $i = 1, \cdots, n$,

$$F_t^i(w) = x_{t+1}^i - f_t^i(x_t, u_t), \qquad F_{T-1}^i(w) = \hat{x}_T^i - f_{T-1}^i(x_{T-1}, u_{T-1}).$$

Sea además,

$$\Xi := \left\{ w = (x_1, \cdots, x_{T-1}, u_0, \cdots, u_{T-1}) \in W \;\middle|\; \begin{array}{c} u_t \in U_t,\ x_{t+1} \in X_{t+1}, \\ \text{para } t = 0, \cdots, T-1 \end{array} \right\}.$$

Con estas definiciones el problema de control discreto $\mathcal{P}_T(\hat{x}_0)$ es equivalente al problema de optimización estática con restricciones

$$\mathcal{P}_T(\hat{x}_0) \quad \equiv \quad \min_{\Omega \cap \Xi} -F^0. \qquad\qquad \mathcal{PE}_{\mathcal{B}T}(\hat{x}_0, \hat{x}_T)$$

Para este problema aplicaremos la condición necesaria de Boltyanskii a $\hat{w}$. En esta dirección calculemos la cúpula de Ξ.

Cúpula de Ξ Calculando similarmente al teorema anterior tenemos que en este caso la cúpula de Ξ en $\hat{w}$ será

$$L = (\mathbb{R}^n)^{T-1} \times L_1 \times \cdots \times L_T.$$

Aplicando la Condición Necesaria de Boltyanskii Aplicando el Teorema B.2.4 de Boltyanskii a nuestro problema tenemos:

(α) Existen números $\mu_0 \leq 0$, $\{\psi_i^t \in \mathbb{R} \ : \ \text{con } 1, \cdots, n; \ t = 0, \cdots, T-1\}$ **no todos nulos.**

(β) Para todo $(\delta x_1, \cdots, \delta x_{T-1}; \delta u_0, \cdots, \delta u_{T-1}) \in L$

$$\begin{aligned} &\textstyle\sum_{s=1}^{T-1} \left\langle \left(-\mu_0 \nabla_{x_s} F^0(\hat{w}) + \sum_{\alpha=1}^{n} \sum_{\gamma=0}^{T-1} \bar{\psi}_\alpha^\gamma \nabla_{x_s} F_\gamma^\alpha(\hat{w}) \right) ; \delta x_s \right\rangle + \\ &\quad + \textstyle\sum_{t=0}^{T-1} \left\langle \left(-\mu_0 \nabla_{u_t} F^0(\hat{w}) + \sum_{\alpha=1}^{n} \sum_{\gamma=0}^{T-1} \bar{\psi}_\alpha^\gamma \nabla_{u_t} F_\gamma^\alpha(\hat{w}) \right) ; \delta u_t \right\rangle \leq 0. \end{aligned}$$

Sistema Adjunto A partir de estas dos enunciados se deduce nuestro corolario. En primer lugar por cálculos análogos a los realizados en el teorema 3.5.1 (PdPB) se tiene desde (β) para $t = 1, \cdots, T-1$,

$$\hat{\psi}_{t-1} = a_0 \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) + \hat{\psi}_t \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t), \qquad (\phi)$$

donde $a_0 := -\mu_0$, $\psi_t := (\psi_t^1, \cdots, \psi_t^n)$.

En vista de que $\hat{x}_T$ es constante, ya no se puede calcular ∇_{x_T}, por lo tanto, ya no se puede asegurar $\psi_{T-1} = -\mu_0 \nabla_{x_T} f_T^0(\hat{x}_T)$ y tampoco $\mu_0 < 0$, (ya no hay condición final).

Principio Débil Realizando cálculos análogos a los del Teorema 3.5.1 considerando vectores del tipo $(\underline{0}_1, \cdots, \underline{0}_T; \underline{0}_1, \cdots, \underline{0}_{t-1}, \delta u_t, \underline{0}_{t+1}, \cdots, \underline{0}_T) \in L$, donde $\delta u_t \in L_t$, en (β), obtenemos para $t = 0, \cdots, T-1$,

$$\left\langle \left(a_0 \nabla_{u_t} f_t^0(x_t, u_t) + \hat{\psi}_t \circ D_{u_t} f_t(x_t, u_t) \right) ; \delta u_t \right\rangle \leq 0.$$

$\blacksquare$

Usando este mismo método y el Teorema de Boltyanskii (B.2.4) podemos estudiar el caso más general, en el cual $x_t \in \partial X_t$ (frontera). Obtenemos el siguiente resultado.

Teorema 3.5.2 (PdPB - Generalizado) *Para* $\mathcal{P}_T(\hat{x}_0)$ *asumamos* $(H1)_B$. *Sea la sucesión de decisiones* $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, *con su proceso* $(\hat{x}_1, \cdots, \hat{x}_T)$, *una* **solución óptima**.

Entonces existen $a_0 \geq 0$ *y vectores* $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, **no todos nulos** *tales que:*

1. *Para cada* $t = 1, \cdots, T-1$, *se cumple:*

Para todo $\delta x_t \in C_t$,

$$\left\langle \left(\nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t) - \hat{\psi}_{t-1} \right) ; \delta x_t \right\rangle \leq 0, \qquad (EA)_{Bg}$$

con condición final: Para todo $\delta x_T \in C_T$

$$\left\langle \left(\nabla_{x_t} (a_0 f_T^0(\hat{x}_T)) - \hat{\psi}_{T-1} \right) ; \delta x_T \right\rangle \leq 0, \qquad (CT)_{Bg}$$

donde C_t *es la cúpula de* $\hat{x}_t$ *respecto a* X_t.

2. *Para cada* $t = 0, \cdots, T-1$, *se cumple:*

Para toda $\delta u_t \in L_t$,

$$\left\langle \nabla_u H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t) \; ; \; \delta u_t \right\rangle \leq 0, \qquad (PdPBg)$$

donde L_t *es la cúpula de* $\hat{u}_t$ *respecto al conjunto* U_t.

Prueba. La demostración sigue los mismos pasos que el teorema 3.5.1 (PdPB) Lo único que cambiará es que se tomarán en cuenta las cúpulas respecto a los conjuntos X_t.

Del Lema 3.5.1 nuestro problema es equivalente a

$$\min_{\Omega \cap \Xi} -F^0. \qquad \mathcal{PE}_{\mathcal{B}T}(\hat{x}_0)$$

Desde esta equivalencia $\hat{w} := (\hat{x}_1, \cdots, \hat{x}_T; \hat{u}_0, \cdots, \hat{u}_{T-1})$ es una solución óptima para este problema. Procedamos análogamente al Teorema 3.5.1.

Cúpula de $\hat{w}$ Calculemos la cúpula de de $\hat{w}$ respecto a Ξ. Siguiendo los razonamientos similares a los del Teorema 3.5.1, en este caso la cúpula será:

$$L = C_1 \times \cdots \times C_T \times L_0 \times \cdots \times L_{T-1}.$$

Aplicando el Teorema de Boltyanskii Del Teorema B.2.4 de optimalidad de Boltyanskii en nuestro problema $\mathcal{PE}_{\mathcal{B}T}(\hat{x}_0)$, para $\hat{w}$ tenemos:

(α) Existen números $\mu_0 \leq 0$, $\{\psi_i^t \in \mathbb{R}\ :\ \text{ con } 1, \cdots, n;\ t = 0, \cdots, T-1\}$ **no todos nulos.**

(β) Para todo $(\delta x_1, \cdots, \delta x_T, \delta u_0, \cdots, \delta u_{T-1}) \in L$ se cumple:

$$\begin{aligned} &\textstyle\sum_{s=1}^{T-1} \left\langle \left(-\mu_0 \nabla_{x_s} F^0(\hat{w}) + \sum_{\alpha=1}^{n} \sum_{\gamma=0}^{T-1} \bar{\psi}_\alpha^\gamma \nabla_{x_s} F_\gamma^\alpha(\hat{w}) \right); \delta x_s \right\rangle + \\ &\quad + \textstyle\sum_{t=0}^{T-1} \left\langle \left(-\mu_0 \nabla_{u_t} F^0(\hat{w}) + \sum_{\alpha=1}^{n} \sum_{\gamma=0}^{T-1} \bar{\psi}_\alpha^\gamma \nabla_{u_t} F_\gamma^\alpha(\hat{w}) \right); \delta u_t \right\rangle \leq 0. \end{aligned}$$

Sistema Adjunto Desde que (β) se cumple para todo vector de L. Consideremos para $t = 1, \cdots, T$,y $\delta x_t \in C_t$, vectores de la forma

$$(\underline{0}_1, \cdots, \underline{0}_{t-1}, \delta x_t, \underline{0}_{t+1}, \cdots, \underline{0}_T; \underline{0}_1, \cdots, \underline{0}_T) \in L.$$

Por lo cual desde (β) tenemos para $t = 1, \cdots, T$,

$$\left\langle \left(-\mu_0 \nabla_{x_t} F^0(\hat{w}) + \sum_{\alpha=1}^{n} \sum_{\gamma=0}^{T-1} \psi_\alpha^\gamma \nabla_{x_t} F_\gamma^\alpha(\hat{w}) \right);\ \delta x_t \right\rangle \leq 0. \qquad (\hat{\phi})$$

Definiendo $a_0 := -\mu_0 \geq 0$, y $\hat{\psi}_t := (\psi_1^t, \cdots, \psi_n^t)$ y realizando cálculos enteramente análogos a los hechos en el Teorema 3.5.1 desde ($\hat{\phi}$) deducímos lo siguiente:

Para $t = 1, \cdots, T-1$, y $\delta x_t \in C_t$,

$$\left\langle -\mu_0 \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) + \psi_t \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t) - \psi_{t-1}\ ;\ \delta x_t \right\rangle \leq 0, \qquad (\phi)$$

para $\delta x_T \in C_T$,

$$\left\langle \left(\nabla_{x_t}(a_0 f_T^0(\hat{x}_T)) - \hat{\psi}_{T-1} \right); \delta x_T \right\rangle \leq 0.$$

Considerando la definición de los Hamiltonianos en (ϕ) deducímos $(EA)_{Bg}$.

Principio Débil En (β), para $\delta u_t \in L_t$, consideremos vectores del tipo

$$(\underline{0}_1, \cdots, \underline{0}_T; \underline{0}_1, \cdots, \underline{0}_{t-1}, \delta u_t, \underline{0}_{t+1}, \cdots, \underline{0}_T) \in L,$$

obtenemos para $t = 0, \cdots, T-1$,

$$\left\langle \left(a_0 \nabla_{u_t} f_t^0(\hat{x}_t, \hat{u}_t) + \hat{\psi}_t \circ D_{u_t} f_t(\hat{x}_t, \hat{u}_t) \right) \ ; \ \delta u_t \right\rangle \leq 0.$$

con cálculos enteramente análogos a los hechos en el Teorema 3.5.1 deducímos (PdPBg). ■

En la siguiente sección estudiaremos una teoría que generaliza aún más los Principios Débiles y muestra que incluso en el caso más abstracto se conserva el Principio Débil de Pontryaguin para los problemas de control discreto en horizonte finito.

3.6 Principio Débil: Clarke

En esta sección generalizaremos en otro sentido el Principio Débil. Mientras que Boltyanskii lo generaliza para cualquier forma de U_t (pues el concepto de cúpula es bastante general, no se requiere que U_t sea cerrado, compacto, etc.), no obstante, requiere que las funciones f_t y f_t^0 sean de clase C^1. Esta hipótesis se puede debilitar en la aproximación de Clarke, sólo se pedirá que sean funciones "Local Lipschitzs" y que los U_t sean cerrados. Con esta aproximación logramos obtener un teorema que usa directamente la condición necesaria en forma primal, (usando el cono Bouligand, que generaliza el concepto de cúpula en la Teoría de Boltyanskii).

En esta sección pediremos las siguientes hipótesis. Para $t = 0, \cdots, T-1$,

- f_t^0 es Local Lipschitzs y regular en $A_t \times B_t$, y f_T^0 lo es en A_T. $(H1)_C$
- f_t es estrictamente diferenciable en $A_t \times B_t$. $(H2)_C$
- El conjunto U_t es cerrado. $(H3)_C$
- Los estados son interiores, es decir, $x_{t+1} \in intX_{t+1}$. $(I)_C$

Equivalencia con un Problema de Optimización Estático Nuevamente la metodología para estudiar esta aproximación será transformar el problema original en uno estático. Esta transformación es la misma que se ha definido en el Lema 3.5.1, debido a que los elementos que definen las funciones f_t^0, f_t (sus dominios y reglas de asignación) siguen siendo los mismos en esta sección.

En conclusión nuestro problema de control óptimo discreto $\mathcal{P}_T(\hat{x}_0)$ es equivalente al problema estático $\mathcal{PE}_{\mathcal{B}}(\hat{x}_0)$. Por esto podemos usar en este problema equivalente la condición necesaria de Clarke.

Teorema 3.6.1 (Principio Débil de Pontryagin-Clarke (PdPC)) *En el problema $\mathcal{P}_T(\hat{x}_0)$ asumamos $(H1)_C$, $(H2)_C$ y $(H3)_C$. Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso $(\hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima***, satisfaciendo $(I)_C$.*

Entonces existen funcionales lineales $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in (\mathbb{R}^n)^*$ **no todas nulas** *tales que:*

1. *Para $t = 1, \cdots, T-1$,*

$$\hat{\psi}_{t-1} \in \partial_{x_t} H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \tag{$(EA)_C$}$$

con condición final $\hat{\psi}_{T-1} \in \partial_{x_T} f_T^0(\hat{x}_T)$.

2. *Para $t = 0, \cdots, T-1$,*

$$\partial_{u_t} H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) \cap N_{U_t}(\hat{u}_t) \neq \emptyset, \tag{$(PdPC)_t$}$$

donde $N_{U_t}(\hat{u}_t)$ es el cono normal de U_t en $\hat{u}_t$.

Prueba. Desde la optimalidad de la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, deducímos que $\hat{w} := (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ es solución óptima del problema $\mathcal{PE}_{\mathcal{B}T}(\hat{x}_0)$.

A diferencia de todos los teoremas anteriores, no podemos aplicar directamente la condición necesaria de Clarke. Para aplicarlo correctamente requerímos que Ξ sea cerrado lo cual no es cierto necesariamente (sólo hemos supuesto que U_t es cerrado, de X_t no se ha dicho nada, sin embargo no es necesario añadir alguna hipótesis adicional).

Por este motivo procederemos de la siguiente manera. Por la hipótesis $(I)_C$, existen bolas cerradas para cada x_t tales que $\bar{B}(\hat{x}_t) \subset \operatorname{int} X_t$. Definamos el conjunto:

$$C = \bar{B}(\hat{x}_1) \times \cdots \times \bar{B}(\hat{x}_T) \times U_0(\hat{u}_0) \times \cdots \times U_{T-1}(\hat{u}_{T-1}).$$

Es claro que $\hat{z} = (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1}) \in C$ por ello $C \neq \emptyset$.

Desde que $C \subset \Xi$ se tiene que: $\hat{z}$ es solución óptima de

$$\min_{\Omega \cap C} -F^0. \tag{ϕ}$$

Definiendo el Lagrangiano Generalizado Definamos el Lagrangiano Generalizado $\mathcal{L} : W \times \mathbb{R} \times (\mathbb{R}^n)^T \times \mathbb{R} \to \mathbb{R}$,

$$\mathcal{L}(w, a_0, (\lambda_t)_t, l) := -a_0 F^0(w) + \sum_{t=0}^{T-1} \langle \lambda_t, F_t^h(w) \rangle + l \|(a_0, (\lambda_t)_t)\| \, d_C(w),$$

donde $(\lambda_t)_t := (\lambda_0, \cdots, \lambda_{T-1}) \in (\mathbb{R}^n)^T$, $\|(a_0, (\lambda_t)_t)\|$, es la norma del vector en el espacio $\mathbb{R} \times (\mathbb{R}^n)^T$, y $d_C(w)$ es la función distacia de w respecto a C.

Para facilitar los cálculos posteriores vamos a simplificar los términos del Lagrangiano Generalizado. Teniendo en cuenta las definiciones de las funciones F_t^h y F^0, las siguientes expresiones son equivalentes al Lagrangiano.

$$-a_0 \left[\sum_{t=0}^{T-1} f_t^0(x_t, u_t) + f_T^0(x_T) \right] + \sum_{t=0}^{T-1} \langle \lambda_t, x_{t+1} - f_t(x_t, u_t) \rangle + l \|(a_0, (\lambda_t)_t)\| d_C(\hat{w}),$$

$$\sum_{t=0}^{T-1} \langle \lambda_t, x_{t+1} \rangle - \sum_{t=0}^{T-1} [a_0 f_t^0(x_t, u_t) + \langle \lambda_t, f_t(x_t, u_t) \rangle] - a_0 f_T^0(x_T) + l \|(a_0, (\lambda_t))\| d_C(\hat{w}),$$

$$\sum_{t=0}^{T-1} \langle \lambda_t, x_{t+1} \rangle \ - \sum_{t=0}^{T-1} H_t(x_t, u_t, a_0, \lambda_t) \ - a_0 f_T^0(x_T) + \ l \ \|(a_0, (\lambda_t)_t)\| \ d_C(\hat{w}),$$

donde $H_t(x_t, u_t, a_0, \lambda_t) := a_0 f_t^0(x_t, u_t) - \langle \lambda_t, f_t(x_t, u_t) \rangle$.

Aplicación de la Condición Necesaria de Clarke Por lo dicho anteriormente $\hat{w} = (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ es solución de (ϕ), entonces del Teorema de Lagrange Generalizado (Teorema A.2.3) tenemos:

Para l suficientemente grande, existen $a_0 \in \mathbb{R}$ y vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$ que cumplen:

(A) $a_0 \geq 0$.

(B) $(a_0, \psi_0, \cdots, \psi_{T-1})$ es **no nulo**.

(C) $0 \in -\partial_{(\vec{x},\vec{u})} \mathcal{L}(\hat{x}, \hat{u}, a_0, \psi_0, \cdots, \psi_{T-1}, l)$.

De un modo análogo a los teoremas anteriores (3.3.1, 3.5.1, etc.) a partir del cálculo de (C) se obteniene el teorema. Por ello usaremos las herramientas del Cálculo Generalizado de Clarke (ver Apéndice A).

Cálculo de Gradientes Generalizados Durante estos cálculos usaremos la conveción de que $x := (x_1, \cdots, x_T) \in (R^n)^T$, $u := (u_0, \cdots, u_{T-1}) \in (R^m)^T$ y $w := (x, u)$, en particular $\hat{w} = (\hat{x}, \hat{u})$, con $\hat{x} := (\hat{x}_1, \cdots, \hat{x}_T)$, y $\hat{u} := (\hat{u}_0, \cdots, \hat{u}_{T-1})$.

Calculemos la derivada generalizada de $\mathcal{L}$, $-\partial_{(x,u)} \mathcal{L}(\hat{w}, a_0, \psi_0, \cdots, \psi_{T-1}, l)$. Sabemos que el Lagrangiano es igual a

$$\sum_{t=0}^{T-1} \langle \psi_t, \hat{x}_{t+1} \rangle \ - \sum_{t=0}^{T-1} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t) \ - a_0 f_T^0(\hat{x}_T) + \ l \ \|(a_0, (\psi_t)_t)\| \ d_C(\hat{w}),$$

por lo que de la Proposición A.2.1. $-\partial_{(x,u)}\mathcal{L}(\hat{w}, a_0, \psi_0, \cdots, \psi_{T-1}, l)$, esta incluído en

$$\partial_{(x,u)}\left[\sum_{t=0}^{T-1}(H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t) - \langle\psi_t, \hat{x}_{t+1}\rangle + a_0 f_T^0(x_T))\right] + \partial_{(x,u)} l \, \|a_0, (\psi_t)\| d_C(\hat{w}).$$

de la misma proposición

$$\partial_{(x,u)}\left[l \, \|a_0, (\psi_t)_t\| \, d_C(\hat{w})\right] = l \, \|a_0, (\psi_t)_t\| \, \partial_{(x,u)} d_C(\hat{w}).$$

De la Proposición A.2.6 tenemos que: $l \, \|a_0, (\psi_t)_t\| \, \partial_{(x,u)} d_C(\hat{w}) \subseteq N_C(\hat{w})$. Por lo que a partir de **(C)**

$$0 \in \partial_{(x,u)}\left[\sum_{t=0}^{T-1}(H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t) - \langle\psi_t, \hat{x}_{t+1}\rangle + a_0 f_T^0(x_T))\right] + N_C(\hat{x}, \hat{u}). \quad (1)$$

Calculemos $N_C(\hat{x}, \hat{u})$, desde que $C = \prod_{t=1}^{T} \bar{B}(\hat{x}_t) \times \prod_{t=1}^{T} U_t(\hat{u}_t)$, usando el Corolario A.2.1 obtenemos

$$N_C(\hat{x}, \hat{u}) = \prod_{t=1}^{T} N_{B_t}(\hat{x}_t) \times \prod_{t=0}^{T-1} N_{U_t}(\hat{u}_t),$$

de otro lado como $\hat{x}_t$ esta en el interior de B_t, $N_{B_t}(\hat{x}_t) = \{0\}$ por ello

$$N_C(\hat{x}, \hat{u}) = \{0\} \times \prod_{t=1}^{T-1} N_{U_t}(\hat{u}_t).$$

A continuación calculemos

$$\partial_{(x,u)}\left[\sum_{t=0}^{T-1}(H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t) - \langle\psi_t, \hat{x}_{t+1}\rangle) + a_0 f_T^0(x_T)\right],$$

para ello probemos en primer lugar que bajo nuestras hipótesis es **regular** la siguiente función (ver definición en Apéndice A)

$$\sum_{t=0}^{T-1}(H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t) - \langle\psi_t, \hat{x}_{t+1}\rangle) + a_0 f_T^0(x_T).$$

En efecto, por hipótesis $f_t(x_t, u_t)$ es estrictamente diferenciable, esto es para toda $(v, w) \in \mathbb{R}^n \times \mathbb{R}^m$,

$$\lim_{\substack{(y,u)\to(\hat{x}_t,\hat{u}_t) \\ s\downarrow 0}} \frac{f_t((y,u) + s(v,w)) - f_t(y,u)}{s} = D_s f_t(\hat{x}_t, \hat{u}_t) \cdot (v, w),$$

en consecuencia la derivada direccional generalizada será

$$\begin{aligned} f_t^o((\hat{x}_t,\hat{u}_t);(v,w)) &= \limsup_{\substack{(y,u)\to(\hat{x}_t,\hat{u}_t)\\ s\downarrow 0}} \tfrac{f_t((y,u)+s(v,w))-f_t(y,u)}{s} \\ &= D_s f_t(\hat{x}_t,\hat{u}_t)\cdot(v,w). \end{aligned}$$

Observando que la función g_t es la composición de la funcional lineal (ψ_t) con la función $f_t(x_t,u_t)$ tenemos para toda $(v,w)\in\mathbb{R}^n\times\mathbb{R}^m$,

$$\begin{aligned} g_t^o((\hat{x}_t,\hat{u}_t);(v,w)) &= \limsup_{\substack{(y,u)\to(\hat{x}_t,\hat{u}_t)\\ s\downarrow 0}} \tfrac{\langle\psi_t, f_t((y,u)+s(v,w))-f_t(y,u)\rangle}{s} \\ &= \langle\psi_t,\ D_s f_t(\hat{x}_t,\hat{u}_t)\cdot(v,w)\rangle. \end{aligned}$$

De otro lado es claro que

$$\lim_{s\downarrow 0} \tfrac{g_t((\hat{x}_t,\hat{u}_t)+s(v,w))-g_t(y,u)}{s} = \lim_{\substack{(y,u)\to(\hat{x}_t,\hat{u}_t)\\ s\downarrow 0}} \tfrac{g_t((y,u)+s(v,w))-g_t(y,u)}{s},$$

en consecuencia para toda $(v,w)\in\mathbb{R}^n\times\mathbb{R}^m$: $g_t'(\hat{x}_t,\hat{u}_t;(v,w)) = g_t^o(\hat{x}_t,\hat{u}_t;(v,w))$, por lo tanto g_t es una función regular (g_t' es la derivada direccional usual).

Por cálculos similares también las funciones $\langle -\psi_t, x_{t+1}\rangle$ y $a_0 f_T^0$ son regulares. Finalmente desde que

$$H_t(x_t,u_t,a_0,\psi_t) - \langle\psi_t,x_{t+1}\rangle = a_0 f_t^0(x_t,u_t) + g_t(x_t,u_t) + \langle -\psi_t, x_{t+1}\rangle,$$

es suma lineal de funciones regulares (con constantes no negativas) usando la Proposición A.2.3 (c) deducimos que H_t es regular. Por la misma razón lo es la función

$$\sum_{t=0}^{T-1}(H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t) - \langle\psi_t,\hat{x}_{t+1}\rangle) + a_0 f_T^0(x_T).$$

Denotemos momentáneamente a esta función como $G(x,u,a_0,\psi)$. Desde su regularidad podemos aplicar la Proposición A.2.5 en $\hat{w}=(\hat{x},\hat{u})$,

$$\partial_{(x,u)}G(\hat{x},\hat{u},a_0,\psi) \subseteq (\partial_x G(\hat{x},\hat{u},a_0,\psi)\ ,\ \partial_u G(\hat{x},\hat{u},a_0,\psi),$$

por esto y los cálculos anteriores, en **(1)** tenemos:

2.1 $0\in -\partial_x\left[\sum_{t=0}^{T-1}(H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t) - \langle\psi_t,\hat{x}_{t+1}\rangle) + a_0 f_T^0(\hat{x}_T)\right] + \{0\}$,

2.2 $0\in -\partial_u\left[\sum_{t=0}^{T-1}(H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t) - \langle\psi_t,\hat{x}_{t+1}\rangle) + a_0 f_T^0(\hat{x}_T)\right] + \prod_{t=0}^{T-1} N_{U_t}(\hat{u}_t)$.

Teniendo en cuenta las variables $(x_1,\cdots,x_t)$ y $(u_1,\cdots,u_t)$ ambas expresiones son equivalentes a

2.1 $0\in\ \partial_{x_k}\left[\sum_{t=0}^{T-1}(H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t) - \langle\psi_t,\hat{x}_{t+1}\rangle) + a_0 f_T^0(\hat{x}_T)\right]$,

2.2 $\qquad 0 \in -\partial_{u_r}\left[\sum_{t=0}^{T-1}\left(H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t)-\langle\psi_t,\hat{x}_{t+1}\rangle\right)+a_0 f_T^0(\hat{x}_T)\right]+N_{U_k}(\hat{u}_k).$

De **2.1**, para $t=1,\cdots,T-1$,

$$0 \in \partial_{x_t}H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t)-\partial_{x_t}(\langle\psi_{t-1},\hat{x}_t\rangle),$$

lo que es igual a

$$\psi_{t-1}\in\partial_{x_t}H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t) \qquad (\alpha)_t$$

De otro lado en **2.1** cuando $t=T$ tenemos que para todo $k=1,\cdots,T-1$, $H_k(x_k,u_k,a_0,\psi_k)$ es constante respecto a x_T, también lo es $\langle\psi_k,x_{k+1}\rangle$ (cuando $k\le T-2$). Por ello calculando ∂_{x_T}, tenemos

$$0\in\partial_{x_T}(a_0 f_T^0(\hat{x}_T))-\psi_{T-1}. \qquad (\alpha)_T$$

A continuación probemos que $a_0>0$, en efecto si no lo fuese desde (B) tendríamos $a_0=0$, por lo cual $a_0 f_T^0(x_T)=0$ entonces $\partial_{x_T}(a_0 f_T^0(\hat{x}_T))=0$. Desde $(\alpha)_T$ deducímos que en este caso $\psi_{T-1}=0$ y por ello

$$0=a_0 f_{T-1}^0(x_{T-1},u_{T-1})+\langle\psi_{T-1},f_{T-1}(x_{T-1},u_{T-1})\rangle=0,$$

usando $(\alpha)_t$ tendríamos que $\psi_{T-2}=0$. Recursivamente obtendríamos

$$0=\psi_0=\cdots=\psi_{T-1},$$

lo cual contradice (A) por lo tanto $a_0>0$.

En vista de que $a_0>0$ definamos para $t=0,\cdots,T-1$,

$$\hat{\psi}_t=\frac{\psi_t}{a_0}.$$

Con estos en $(\alpha)_t$ tenemos para $t=1,\cdots,T-1$,

$$\hat{\psi}_{t-1}\in\partial_{x_t}H_t(\hat{x}_t,\hat{u}_t,\hat{\psi}_t)$$

y en $(\alpha)_T$ tenemos $\hat{\psi}_{T-1}\in\partial_{x_T}f_T^0(\hat{x}_T)$.

Por lo que hemos demostrado que estas funcionales $(\hat{\psi}_t)$ satisfacen $(EA)_C$.

Finalmente de los cálculos respecto a las variables u_t tenemos desde **2.2** que para $t=0,\cdots,T-1$,

$$0\in-\partial_{u_t}H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t)+N_{U_t}(\hat{u}_t),$$

lo que equivale a

$$\partial_{u_t}H_t(\hat{x}_t,\hat{u}_t,a_0,\psi_t)\cap N_{U_t}(\hat{u}_t) \qquad (\phi).$$

De la definición de $\hat{\psi}_t$, es claro que $\partial_{u_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t) = a_0 \partial_{u_t} H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t)$. Teniendo en cuenta además que $N_{U_t}(\hat{u}_t)$ es un cono, no es difícil deducir a partir de (ϕ) que para cada $t = 0, \cdots, T-1$,

$$\nabla_u H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) \cap N_{U_t}(\hat{u}_t) \neq \emptyset.$$

■

Observación. Este resultado generaliza el Principio Débil de Pontryagin-Boltyanskii (Teorema 3.5.1), por dos razones: 1) La condición

$$\nabla_u H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) \cap N_{U_t}(\hat{u}_t) \neq \emptyset,$$

es equivalente a: Para toda $\delta u_t \in B_{U_t}(\hat{u}_t)$,

$$\langle \nabla_u H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \delta u_t \rangle \leq 0,$$

donde $B_{U_t}(\hat{u}_t)$ es el Cono Bouligand de $\hat{u}_t$ respecto a U_t.

Esta desigualdad es similar a la dada en Teorema 3.5.1, en la cual sólo se usa cúpulas. Sabemos que toda cúpula L_t de $\hat{u}_t$ respecto a U_t está incluído en el cono Bouligand, por lo que este resultado es más general.

2) En este Teorema se usa funciones más generales (Local Lipschitz) que en el caso de Boltyanskii (funciones de tipo C^1).

Estudiemos algunas consecuencias del Teorema 3.6.1. Siguiendo la misma metodología de demostración, podemos abordar también el problema de control discreto con extremo derecho fijo (con objetivo final) $\hat{x}_T$, es decir el problema $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$.

Corolario 3.6.1 (PdPC - Extremo Fijo) *Para el problema $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$ asumamos $(H1)_C$, $(H2)_C$ y $(H3)_C$. Sea $(\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ una* **solución** *del problema, satisfaciendo $(I)_C$.*

Entonces existen $a_0 \geq 0$ y funcionales lineales $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in (\mathbb{R}^n)^*$ **no todo nulo** *tales que:*

1. *Para $t = 1, \cdots, T-1$,*

$$\hat{\psi}_{t-1} \in \partial_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t). \qquad (EA)_C$$

2. *Para $t = 0, \cdots, T-1$,*

$$\partial_{u_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t) \cap N_{U_t}(\hat{u}_t) \neq \emptyset, \qquad (PdPC)_t$$

donde $N_{U_t}(\hat{u}_t)$ es el cono normal de U_t en $\hat{u}_t$.

Con $H_t(x_t, u_t, a_0, \psi_t) = a_0 f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$ el hamiltoniano.

Prueba. La demostración de este corolario es similar a la del Corolario 3.5.1, que es el análogo de éste, para el Principio Débil de Pontryagin-Boltyanskii. Como en dicha demostración la única diferencia que debemos considerar aquí es que la variable x_T es constante ($x_T = \hat{x}_T$), por lo que en nuestros cálculos tenemos una variable menos, en consecuencia ya no se podrá obtener la condición final para el sistema adjunto $(EA)_B$ y tampoco se puede demostrar que $a_0 \neq 0$.

A parte de esto todos los cálculos son los mismos, por lo que no es difícil demostrar este Corolario. ■

Este mismo Corolario podemos expresarlo de un modo equivalente como sigue.

Corolario 3.6.2 *Para el problema $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$ asumamos $(H1)_C - (H3)_C$. Sea $(\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ una* **solución** *del problema, satisfaciendo $(I)_C$.*

Entonces existen $a_0 \geq 0$ y funcionales lineales $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in (\mathbb{R}^n)^*$ **no todo nulo** *tales que:*

1. *Para cada $t = 1, \cdots, T-1$, existe $\varphi_t \in \partial_{x_t} f_t^0(\hat{x}_t, \hat{u}_t)$*

$$\hat{\psi}_{t-1} = a_0 \varphi_t + \hat{\psi}_t \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t). \qquad \widetilde{(EA)}_C$$

2. *Para cada $t = 0, \cdots, T-1$, existe $\phi_t \in \partial_{x_t} f_t^0(\hat{x}_t, \hat{u}_t)$, tal que se satisface:*

Para todo $\delta w_t \in B_{U_t}(\hat{u}_t)$

$$\langle a_0 \phi_t + \hat{\psi}_t \circ D_{u_t} f_t(\hat{x}_t, \hat{u}_t), \delta w_t \rangle \leq 0, \qquad \widetilde{(PdPC)}_t$$

donde $B_{U_t}(\hat{u}_t)$ es el cono Bouligand de U_t en $\hat{u}_t$.

Prueba. Para probar este corolario solamente debemos hacer algunos cálculos en el corolario anterior.

Con nuestras hipótesis, del cálculo generalizado, para $t = 1, \cdots, T-1$,

$$\begin{aligned}
\partial_{x_t} H_t(x_t, u_t, a_0, \psi_t) &= \partial_{x_t}[a_0 f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle] \\
&= \partial_{x_t}[a_0 f_t^0(x_t, u_t) + \textstyle\sum_{i=1}^n \psi_t^i f_t^i(x_t, u_t)] \\
&= a_0 \partial_{x_t} f_t^0(x_t, u_t) + \textstyle\sum_{i=1}^n \psi_t^i \partial_{x_t} f_t^i(x_t, u_t)
\end{aligned}$$

Desde que f_t^i son estrictamente diferenciables,

$$\partial_{x_t} f_t^i(x_t, u_t) = D_{x_t} f_t^i(x_t, u_t)$$

En consecuencia tenemos

$$\partial_{x_t} H_t(x_t, u_t, a_0, \psi_t) = a_0\, \partial_{x_t} f_t^0(x_t, u_t) + \psi_t \circ D_{x_t} f_t^i(x_t, u_t).$$

De la definición de la gradiente generalizada de $\partial_{x_t} H_t$, el sistema adjunto $(EA)_C$ es equivalente a $(\widetilde{EA_C})$.

Por consideraciones similares para el cálculo de $\partial_{u_t} H_t(x_t, u_t, a_0, \psi_t)$ y desde que el cono normal $N_{U_t}(\hat{u}_t)$ es ortogonal al cono Bouligand, deducímos que $(\widetilde{PdPB})_t$ es equivalente a $(PdPB)_t$. ■

Con similares consideraciones para el Teorema 3.6.1 y asumiendo, en vez de $(H1)_C$, una hipótesis más fuerte

- Para $t = 0, \cdots, T$,

$$f_t^0 \quad \text{es estrictamente regular en su dominio.} \qquad (\widetilde{H1})_C$$

obtenemos el siguiente teorema que generaliza el Principio Débil de Pontryaguin-Boltyanskii (Teorema 3.5.1)

Teorema 3.6.2 (Generalización del PdPB) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ asumamos $(\widetilde{H1})_C$, $(H2)_C$ y $(H3)_C$ Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso $(\hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima** *que satisfaze $(I)_C$.*

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, **no todos nulos** *tales que:*

1. *Para $t = 1, \cdots, N-1$,*

$$\hat{\psi}_{t-1} = \nabla_x H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \qquad (EA)$$

con condición final, $\hat{\psi}_{N-1} = \nabla_x f_T^0(\hat{x}_T)$.

2. *Para $t = 0, \cdots, N-1$, se cumple: Para cualquier $\delta u \in B_{U_t}(\hat{u}_t)$,*

$$\langle \nabla_u H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) \; ; \; \delta u \rangle \leq 0, \qquad (PdPB)_t$$

donde $B_{U_t}(\hat{u}_t)$ es el cono Bouligand de $\hat{u}_t$.

Prueba. Desde la hipótesis $(\widetilde{H1})_C$ tenemos $\partial_{x_t} f_t(\hat{x}_t, \hat{u}_t) = \{\nabla_{x_t} f_t(\hat{x}_t, \hat{u}_t)\}$, por lo que

$$\partial_{x_t} H(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) = \{\nabla_{x_t} H(\hat{x}_t, \hat{u}_t, \hat{\psi}_t)\}.$$

A partir de esto y de $(EA)_C$ se deduce (EA).

Finalmente como el cono normal es ortogonal al cono Bouligand, deducímos de (EA) y $(PdPC)_t$ la parte **2.** del teorema actual. ■

Con estas mismas hipótesis el Corolario 3.6.2, toma la siguiente forma

Corolario 3.6.3 *En el problema $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$ asumamos $(\widetilde{H1})_C$, $(H2)_C$ y $(H3)_C$. Sea $(\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ una* **solución** *del problema, satisfaciendo $(I)_C$.*

Entonces existen $a_0 \geq 0$ y funcionales lineales $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in (\mathbb{R}^n)^*$ **no todo nulo** *tales que:*

1. *Para cada* $t = 1, \cdots, T-1$,

$$\hat{\psi}_{t-1} = \nabla_x H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t). \qquad (\widetilde{EA})_C$$

2. *Para cada* $t = 0, \cdots, T-1$, *existe* $\phi_t \in \partial_{x_t} f_t^0(\hat{x}_t, \hat{u}_t)$, *tal que se satisface:*

Para todo $\delta w_t \in B_{U_t}(\hat{u}_t)$

$$\langle \nabla_x H_t(\hat{x}_t, \hat{u}_t, a_0, \hat{\psi}_t)\ ,\ \delta w_t \rangle \leq 0, \qquad (\widetilde{PdPC})_t$$

donde $B_{U_t}(\hat{u}_t)$ *es el cono Bouligand de* U_t *en* $\hat{u}_t$.

Prueba. Desde la hipótesis $(\widetilde{H}1)_C$ tenemos

$$\partial_{x_t} H(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) = \{\nabla_{x_t} H(\hat{x}_t, \hat{u}_t, \hat{\psi}_t)\}.$$

Por lo que del Corolario 3.6.2, se deduce el presente. ■

Observación. Todos los principios débiles demostrados hasta aquí, nos muestran que en general las decisiones $\hat{u}_t$, son puntos críticos para subproblemas $(PM)_t$. ¿Serán incluso, soluciones óptimas? La respuesta es **no.** Esto será estudiado en la siguiente sección.

3.7 Contraejemplo de Boltyanskii

Hemos visto que en general se cumplía que cada control óptimo $\hat{u}_0, \cdots, \hat{u}_{T-1}$, era punto crítico del subproblema

$$\max_{u \in U_t} H_t(\hat{x}_t, u, \hat{\psi}_t). \qquad (PM)_t$$

(Observemos que incluso cuando el hamiltoniano tiene la forma más general $H_t(\hat{x}_t, u, \hat{\psi}_t, \nu_t)$ (Teorema 3.4.1) sigue siendo cierto que $\hat{u}_t$ es un punto crítico del subproblema $(PM)_t$).

Es natural preguntarse dentro de la equivalencia que podría haber entre el problema de control continuo y el discreto, si es posible que cada control $\hat{u}_t$ sea de hecho una **solución óptima** del subproblema $(PM)_t$ (esto es el PFP).

El matemático V.G. Boltyanskii encontró un contraejemplo en el cual la solución óptima no satisface el Principio Fuerte del Máximo.

Pasemos a proponer 2 contraejemplos de Boltyanskii. En el primero se demuestra que el PFP no es una condición necesaria, y en el segundo se demuestra que tampoco es una condición suficiente.

Contraejemplo: Condición Necesaria En el problema $\mathcal{P}_T(\hat{x}_0)$ tomemos $T \in \mathbb{N}$, fijo y arbitrario.

Definamos los conjuntos: $X_t := \mathbb{R}$ y $U_t := [-1, 1]$. Para $t = 0, \cdots, T-1$, las funciones de evolución y de beneficio serán:

$$f_t := f : \begin{array}{ccc} \mathbb{R} \times [-1,1] & \to & \mathbb{R}, \\ (x,u) & \mapsto & u, \end{array} \qquad f_t^0 := f : \begin{array}{ccc} \mathbb{R} \times [-1,1] & \to & \mathbb{R}, \\ (x,u) & \mapsto & u^2 - 2x^2. \end{array}$$

La función final $f_T \equiv 0$, y nuestro punto inicial $\hat{x}_0 = 0$.

Con estas definiciones el problema $\mathcal{P}_T(\hat{x}_0)$ será.

$$\begin{array}{ll} \max & B_T(x_0, u_0, \cdots, u_{T-1}) := \sum_{t=0}^{T-1} (u_t^2 - 2x_t^2) \\ \text{s.a.} & \qquad \text{para todo } t = 0, \cdots, T-1, \\ & \qquad x_{t+1} \in \mathbb{R},\ u_t \in [-1, 1], \\ & \qquad x_{t+1} = u_t,\ x_0 = 0 \in \mathbb{R}. \end{array} \qquad (\mathcal{P}_T(\hat{x}_0))$$

Para este problema calculemos los Hamiltonianos y el Sistema Adjunto.

En primer lugar el Hamiltoniano será

$$H_t(x_t, u_t, \psi_t) = u_t^2 - 2x_t^2 + \psi_t\, u_t,$$

El Sistema Adjunto será:
Para $t = 1, \cdots, T-1$,

$$\psi_{t-1} = \frac{\partial H_t}{\partial x_t}(x_t, u_t, \psi_t) = -4\, x_t. \qquad (EA)$$

con condición final $\psi_{T-1} = 0$.

Solución Óptima La sucesión $\hat{u}_0 = \hat{u}_1 = \cdots = \hat{u}_{T-2} = 0$ y $\hat{u}_{T-1} = 1$, con su proceso asociado $\hat{x}_0 = \hat{x}_1 = \cdots = \hat{x}_{T-1} = 0$, $\hat{x}_T = 1$, **es una solución óptima**.

En efecto por la forma de las funciones tenemos en general para una sucesión $(u_0, \cdots, u_{T-1}) \in \operatorname{Adm} \mathcal{P}_T(0)$

$$\begin{array}{rcl} B_T(0, u_0, \cdots, u_{T-1}) & = & (u_0^2 - 2(x_0)^2) + \cdots + (u_{T-1}^2 - 2(x_{T-1})^2) \\ & = & u_0^2 + \cdots + u_{T-1}^2 - 2x_1^2 - \cdots - 2x_{T-1}^2 \\ & = & u_{T-1}^2 - u_0^2 - \cdots - u_{T-2}^2 \le u_{T-1}^2 \le 1. \end{array} \qquad (\phi)$$

En particular para nuestro camino $(\hat{u}_0, \hat{u}_1, \cdots, \hat{u}_{T-1})$, obtenemos

$$B_T(0, \hat{u}_0, \hat{u}_1, \cdots, \hat{u}_{T-1}) = 1,$$

por lo tanto es una solución óptima, más aún, por la forma de B_T en (ϕ) esta solución es única.

Principio Fuerte de Pontryagin (PFP) Veamos si esta solución óptima satisface el PFP. Del sistema (EA) hallamos sus variables adjuntas.

$$0 = \hat{\psi}_{T-1} = \cdots = \hat{\psi}_0.$$

De lo cual, para $t = 0, \cdots, T-1$, y $u_t \in U_t$,

$$H_t(\hat{x}_t, u_t, \hat{\psi}_t) = u_t^2 - 2\hat{x}_t^2 - 0 \cdot u_t = u_t^2 - 2(0)^2 = u_t^2.$$

Finalmente si esta solución óptima satisface el PFP (Proposición 3.1.2) entonces tendríamos

Para $t = 0, \cdots, T-1$, $\hat{u}_t = 0$ es solución del subproblema

$$\max_{u_t \in U_t} H(\hat{x}_t, u_t, \hat{\psi}_t) = \max_{u_t \in [-1,1]} u_t^2.$$

Es claro que $\hat{u}_t = 0$ no es una solución. Luego, no se cumple el Principio de Fuerte de Pontryagin (PFP).
Observación. El contraejemplo anterior usa funciones de tipo C^∞, de hecho las funciones f_t son afines, y las funciones f_t^0 son convexas respecto a u y cóncavas para x, por esto seguirá siendo un contraejemplo dentro de varios contextos y condiciones.

Contraejemplo: Condición Suficiente El Principio Fuerte del Máximo tampoco es una condición suficiente. Es decir una sucesión $(u_0, \cdots, u_{T-1})$ con sus proceso asociado $(x_0, \cdots, x_T)$ que satisface el PFP (i.e. para $t = 0, \cdots, T-1$, u_t es solución del problema $(PM)_t$, donde $(\psi_0, \cdots, \psi_{T-1})$ son soluciones de (EA)) no tiene porque ser solución óptima.

A este respecto definamos el siguiente problema $\mathcal{P}_T(\hat{x}_0)$,

$$\begin{array}{ll} \max & B_T(x_0, u_0, \cdots, u_{T-1}) := \sum_{t=0}^{T-1}(2x_t^2 - u_t^2) \\ \text{s.a.} & \quad\text{para todo } \ t = 0, \cdots, T-1, \\ & \quad x_{t+1} \in \mathbb{R}, \ u_t \in [-1,1], \\ & \quad x_{t+1} = u_t, \ x_0 = 0 \in \mathbb{R}. \end{array} \tag{$\mathcal{P}_T(\hat{x}_0)$}$$

Una sucesión que satisface PFP Para este problema la siguiente sucesión $\hat{u}_0 = \cdots = \hat{u}_{T-1} = 0$, con su proceso $\hat{x}_0 = \cdots = \hat{x}_T = 0$, es una sucesión que satisface el PFP.

En efecto, el Hamiltoniano será $H_t(x_t, u_t, \psi_t) = 2x_t^2 - u_t^2 + \psi_t u_t$, y el Sistema Adjunto será:
Para $t = 1, \cdots, T-1$,

$$\psi_{t-1} = \frac{\partial H_t}{\partial x_t}(x_t, u_t, \psi_t) = 4\,x_t. \tag{EA}$$

con condición final $\psi_{T-1} = 0$.

Entonces para $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ las variables adjuntas serán

$$\hat{\psi}_0 = \cdots = \hat{\psi}_{T-1} = 0.$$

Por lo cual los subproblemas $(PM)_t$ son los siguientes

$$\max_{u \in U_t} H(\hat{x}_t, u_t, \hat{\psi}_t) = \max_{u \in [-1,1]} (2(0)^2 - u^2 + 0 \cdot u) = \max_{u_t \in [-1,1]} -u^2.$$

es claro que $\hat{u}_t = 0$ es solución de este subroblema. Así pues la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ **satisface el PFP.**

Finalmente esta sucesión **no es una solución óptima.** En efecto, calculando el beneficio total de $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ tenemos

$$B_T(0, \hat{u}_0, \cdots, \hat{u}_{T-1}) = 0.$$

Por su parte para la sucesión $\bar{u}_0 = 1, \bar{u}_1 = \cdots = \bar{u}_{T-1} = 0$,

$$B_T(0, \bar{u}_0, \cdots, \bar{u}_{T-1}) = 1$$

Por ello $(\hat{u}_1, \cdots, \hat{u}_{T-1})$ **no es solución óptima**, pero si cumple el PFP.

Capítulo 4

Principio Fuerte de Pontryagin Discreto

En el presente capítulo estudiaremos desde diferentes perspectivas, los requisitos que debe satisfacer un problema para que el PFP sea una condición necesaria de óptimo. Gran parte de esta investigación es nueva (salvo la programación mixta), constituyendo un aporte a la Teoría de Control de Procesos Discretos.

Los contraejemplos de Boltyanskii muestran que en general el PFP no es una condición necesaria de optimalidad. Por contraparte en el Corolario 3.4.3 hemos visto que si lo es. Así pues vemos que si el problema tiene ciertas características especiales el PFP es una condición necesaria de optimalidad.

La motivación para dicho estudio es especificar y señalar con mayor precisión cuando el Principio del Máximo de Pontryagin continua siendo válido al ser discretizado. Este problema no sólo tiene un interés teórico sino también práctico. En efecto, hemos visto que una sucesión de decisiones óptimas $\hat{u}_0, \cdots, \hat{u}_{T-1}$, cumplen en general el principio débil, por ello para $t = 0, \cdots, T-1$, $\hat{u}_t$ es un punto crítico del problema

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t). \qquad (PM)_t$$

Si bien esto nos da una forma de aproximarnos a la solución, el conjunto de puntos críticos puede ser grande. Sin embargo si logramos precisar que $\hat{u}_t$ es una solución de cada subproblema entonces reducimos el número de candidatos (algunas veces considerablemente). Por ello cuando es posible usar el PFP, es mucho más eficiente que los principios débiles para hallar la solución.

Hemos encontrado tres maneras diferentes de dar condiciones al problema $P_T(\hat{x}_0)$ para que el PFP sea una condición necesaria de optimalidad. Estas son: 1) Afinidad en la Variable de Estado "x", 2) Programación Mixta, y 3) Condición Suficiente de KKT.

4.1 Afinidad en la Variable de Estado

En esta sección demostraremos que el PFP es una condición necesaria de optimalidad para el problema $P_T(\hat{x}_0)$, cuando las funciones de beneficio y evolución son afines respecto a la variable "x", más cierta hipótesis adicional.

La importancia de este resultado es mayor debido a que en su demostración, hemos encontrado una relación interesante con la Programación Dinámica y nos permite enfocar el PFP como una "consecuencia " del Principio de Optimalidad de Bellman.

Antes de continuar con los resultados más importantes de esta sección, para evitar inconsistencias vamos a suponer la siguiente hipótesis de trabajo.

Para $t = 0, \cdots, T-1$,

$$f_t(X_t \times U_t) \subseteq X_{t+1}. \tag{$Inclu$}$$

Como consecuencia de esta hipótesis se tiene el siguiente lema.

Lema 4.1.1 *Dado $\hat{x}_t \in X_t$ fijo arbitrario. Se cumple que para toda sucesión de decisiones $(u_t, \cdots, u_{T-1} \in U_t \times \cdots \times U_{T-1})$ existe un proceso $(x_t, \cdots, x_T)$ tal que $(x_t, \cdots, x_T; u_t, \cdots, u_{T-1}) \in \mathrm{Adm}_{\mathrm{T}}^{\mathrm{t}}(\hat{\mathrm{x}}_{\mathrm{t}})$.*

Prueba. Basta definir recursivamente

$$\begin{array}{lcl} x_t & := & \hat{x}_t, \\ x_{t+1} & := & f_t(\hat{x}_t, u_t), \\ \cdots & \cdots & \cdots \\ x_{r+1} & := & f_r(x_r, u_r). \end{array}$$

Esto es posible porque siempre se tendrá desde la hipótesis que $x_{r+1} \in X_{r+1}$. ∎

Si bien el lema es trivial es necesario ponerlo claramente pues depende de este resultado todas las conclusiones siguientes. Recordemos además de Lema 2.1.1 que dado $\hat{x}_t \in X_t$ y una sucesión de decisiones $(u_t, \cdots, u_T)$ define recursivamente un único proceso asociado $(x_t, \cdots, x_T)$. Por ello resumimos la notación

$$B_T^t(x_t, x_{t+1}, \cdots, x_T, u_t, \cdots, u_{T-1}) = B_T^t(x_t, u_t, \cdots, u_{T-1}).$$

Finalmente resaltemos que a lo largo de esta sección las funciones de beneficios f_t^0 y de evolución f_t del problema satisfacerán las siguientes hipótesis

- Están definidas en conjuntos del tipo $A_t \times U_t$, donde A_t es un abierto que contiene a X_t. No se exige ninguna condición sobre el conjunto U_t.

- Son diferenciables respecto a la variable "x".

4.1.1 Caracterización de las Variables Adjuntas

Bajo la hipótesis (*inclu*) para el problema $\mathcal{P}_T(\hat{x}_0)$ es posible definir correctamente lo siguiente.

Asociadas a una sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ definimos las siguientes funciones. Para $t = 1, \cdots, T-1$,

$$\begin{array}{rccl} g_t: & X_t & \to & \mathbb{R} \\ & x & \mapsto & B_T^t(x; \hat{u}_t, \cdots, \hat{u}_{T-1}), \end{array}$$

además definamos la función $g_T \equiv f_T^0$.

Bajo la hipótesis de que f_t^0 y f_t son diferenciables respecto a x se deduce que las funciones g_t son diferenciables.

Proposición 4.1.1 (Cálculo de las Variables Adjuntas) *Consideremos el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo la hipótesis $(Inclu)$ y las señaladas al principio de esta sección (dominio de f_t y f_t^0 y diferenciabilidad respecto a "x"). Dada una suceción de decisiones* **fija y arbitraria** $(\hat{u}_0, \cdots, \hat{u}_{T-1}) \in \mathrm{Adm}_{\mathrm{T}}(\hat{\mathrm{x}}_0)$, *con su proceso asociado $(\hat{x}_0, \cdots, \hat{x}_T)$.*

Sea $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1}$ la única solución del sistema (EA).
Para $t = 1, \cdots, N-1$,

$$\hat{\psi}_{t-1} = \nabla_x H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \qquad (EA)$$

con condición final $\hat{\psi}_{N-1} = \nabla_x f_T^0(\hat{x}_T)$.

Entonces considerando las funciones $g_0, \cdots, g_{T-1}$, asociadas a $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, se tiene para $t = 0, \cdots, T-1$,

$$\hat{\psi}_t = \nabla_x g_{t+1}(\hat{x}_{t+1}).$$

Prueba. La demostración del teorema sale directamente al calcular las gradientes de las funciones g_t. Realizemos estos cálculos.

Desde la definición de la función g_T y por la condición final del sistema (EA) tenemos, $\hat{\psi}_{T-1} = \nabla_x g_T(\hat{x}_T)$. Ahora calculemos la gradiente de g_{T-1}, por su definición,

$$\begin{aligned} \nabla_x g_{T-1}(x) &= \nabla_x (f_T^0(x, \hat{u}_{T-1}) + f_T^0(f_{T-1}(x, \hat{u}_{T-1}))) \\ &= \nabla_x (f_T^0(x, \hat{u}_{T-1}) + \nabla_x f_T^0(f_{T-1}(x, \hat{u}_{T-1})) \circ D_x f_T(x, \hat{u}_{T-1}). \end{aligned}$$

Para $x = \hat{x}_{T-1}$, $f_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1}) = \hat{x}_T$ y $\nabla_x f_T^0(\hat{x}_T) = \hat{\psi}_{T-1}$. Por lo cual tenemos

$$\begin{aligned}
\nabla_x g_{T-1}(\hat{x}_{T-1}) &= \nabla_x(f_T^0(\hat{x}_{T-1}, \hat{u}_{T-1}) + \hat{\psi}_{T-1} \circ D_x f_T(\hat{x}_{T-1}, \hat{u}_{T-1})), \\
&= \nabla_x H_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1}, \hat{\psi}_{T-1}).
\end{aligned}$$

Por lo tanto del sistema (EA) deducímos

$$\hat{\psi}_{T-2} = \nabla_x g_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1}, \hat{\psi}_{T-1}).$$

Supongamos que por inducción hayamos probado que hasta $r+1$ se cumple

$$\hat{\psi}_r = \nabla_x g_{r+1}(\hat{x}_{r+1}).$$

Probemos que para r también es cierto. Por su definición las funciones g_t cumplen para todo $t = 1, \cdots, T$,

$$g_t(x) = f_t^0(x, \hat{u}_t) + g_{t+1}(f_t(x, \hat{u}_t)).$$

Por lo tanto para $t = r$,

$$\nabla_x g_r(x) = \nabla_x f_r^0(x, \hat{u}_r) + \nabla_x g_{r+1}(f_r(x, \hat{u}_r)) \circ Df_r(x, \hat{u}_r),$$

$$\begin{aligned}
\nabla_x g_r(\hat{x}_r) &= \nabla_x f_r^0(\hat{x}_r, \hat{u}_r) + \nabla_x g_{r+1}(f_r(\hat{x}_r, \hat{u}_r)) \circ Df_r(\hat{x}_r, \hat{u}_r) \\
&= \nabla_x f_r^0(\hat{x}_r, \hat{u}_r) + \nabla_x g_{r+1}(\hat{x}_{r+1}) \circ Df_r(\hat{x}_r, \hat{u}_r) \\
&= \nabla_x f_r^0(\hat{x}_r, \hat{u}_r) + \hat{\psi}_r \circ Df_r(\hat{x}_r, \hat{u}_r) \\
&= \nabla_x H_r(\hat{x}_r, \hat{u}_r, \hat{\psi}_r).
\end{aligned}$$

Del sistema (EA) obtenemos

$$\psi_{r-1} = \nabla_x g_r(\hat{x}_r).$$

■

Continuando con nuestra investigación definamos las siguientes funciones. Consideremos una sucesión de decisiones fija y arbitraria $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, para $t = 0, \cdots, T-1$, definamos las funciones $\rho_t : X_t \to \mathbb{R}$,

$$\rho_t(x_t) := \sup_{u_t \in U_t} \left\{ f_t^0(x_t, u_t) + B_T^{t+1}(f_t(x_t, u_t); \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) \right\}.$$

Con estas funciones obtendremos una condición necesaria de optimalidad que tiene similitud, con el Corolario 2.2.1 que caracteriza las soluciones usando las funciones valor V_t, y también tiene ciertos aspectos que nos hacen recordar al PFP ($\hat{u}_t$ es solución óptima de cierto subproblema).

Proposición 4.1.2 (Condición Necesaria) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo (inclu), sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso asociado $(\hat{x}_0, \hat{x}_1, \cdots, \hat{x}_T)$ una* **solución óptima**.

Entonces para $t = 0, \cdots, T-1$, $\hat{u}_t$ es **solución óptima** *del subproblema*

$$\max_{u \in U_t} \left\{ f_t^0(\hat{x}_t, u) + B_T^{t+1}(f_t(\hat{x}_t, u); \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) \right\}.$$

Por lo tanto, considerando las funciones ρ_t asociadas a $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, tenemos para $t = 0, \cdots, T-1$,

$$\rho_t(\hat{x}_t) = f_t^0(\hat{x}_t, \hat{u}_t) + B_T^{t+1}(\hat{x}_{t+1}; \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}).$$

Más aún,

$$\rho_t(\hat{x}_t) = V_t(\hat{x}_t),$$

donde V_t son las funciones valor.

Prueba. Tomemos $t = 0, \cdots, T-1$, fijo y arbitrario. Por la definición de la función ρ_t, y como $\hat{u}_t \in U_t$, se sigue que:

$$\begin{array}{rcl} \rho_t(\hat{x}_t) & \geq & f_t^0(\hat{x}_t, \hat{u}_t) + B_T^{t+1}(\hat{x}_{t+1}; \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) \\ & = & B_T^t(\hat{x}_t; \hat{u}_t, \cdots, \hat{u}_{T-1}). \end{array}$$

Desde la optimalidad de las decisiones y usando el Corolario 2.2.1 tenemos $V_t(\hat{x}_t) = B_T^t(\hat{x}_t; \hat{u}_t, \cdots, \hat{u}_{T-1})$. Por lo tanto $\rho_t(\hat{x}_t) \geq V_t(\hat{x}_t)$.

De otro lado usando la definición de la función valor y la Proposición 2.2.1 (ecuación de Bellman)

$$\begin{array}{rcl} \rho_t(\hat{x}_t) & = & \sup_{u_t \in U_t} \{f_t^0(\hat{x}_t, u_t) + B_T^{t+1}(f_t(\hat{x}_t, u_t); \hat{u}_{t+1}, \cdots, \hat{u}_{T-1})\} \\ & \leq & \sup_{u_t \in U_t} \{f_t^0(\hat{x}_t, u_t) + V_{t+1}(f_t(\hat{x}_t, u_t))\} \\ & = & V_t(\hat{x}_t), \end{array}$$

por lo cual

$$\begin{array}{rcl} \rho_t(\hat{x}_t) & = & V_t(\hat{x}_t) \\ & = & B_T^t(\hat{x}_t; \hat{u}_t, \cdots, \hat{u}_{T-1}) \\ & = & f_t^0(\hat{x}_t, \hat{u}_t) + B_T^{t+1}(\hat{x}_{t+1}; \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}). \end{array}$$

■

Observación. Notemos que este resultado es bastante general, sólo requiere de la hipótesis (*inclu*) para su correcta demostración y no necesita de hipótesis adicionales para las funciones de beneficios (f_t^0) y de evolución (f_t).

Con estas proposiciones podemos probar el Principio Fuerte de Pontryagin para problemas cuyas funciones sean afines respecto a la variable "x", esto es problemas que satisfacen la siguiente hipótesis

- Para cada $t = 0, \cdots, T-1$, sea $u_t \in U_t$ fijo y arbitrario, entonces

$$\begin{array}{rl} f_t^0(\cdot, u_t): & X_t \to \mathbb{R}, \\ f_t(\cdot, u_t): & X_t \to \mathbb{R}^n, \end{array} \qquad (H)_A$$

son funciones afines respecto a la variable de estado.

4.1.2 Principio Fuerte de Pontryagin Afín

Teorema 4.1.1 (Principio Fuerte de Pontryagin - Afín (PFP-Afin)) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ asumamos las hipótesis de la Proposición 4.1.1 y la hipótesis $(H)_A$. Sea la sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, con su proceso asociado $(\hat{x}_0, \hat{x}_1, \cdots, \hat{x}_T)$, una* **solución óptima**.

Entonces existen vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$ tales que:

1. *Para $t = 1, \cdots, N-1$,*

$$\hat{\psi}_{t-1} = \nabla_x (H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t), \qquad (EA)$$

con condición final, $\hat{\psi}_{N-1} = \nabla_x f_T^0(\hat{x}_T)$.

2. *Para $t = 0, \cdots, T-1$, $\hat{u}_t$ es* **solución** *del subproblema*

$$\max_{u \in U_t} H_t(\hat{x}, u, \hat{\psi}_t), \qquad (PM)_t$$

donde $H_t(x_t, u_t, \psi_t) = f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$ es el hamiltoniano.

Prueba. Desde la optimalidad de $(\hat{u}_0, \cdots, \hat{u}_{T-1})$,

$$V_0(\hat{x}_0) < \infty \;\Rightarrow\; V_t(\hat{x}_t) < \infty.$$

De la proposición anterior $\rho_t(\hat{x}_t) = V_k(\hat{x}_t)$. Por otra parte consideremos las funciones g_t asociadas a $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, por sus definiciones y la optimalidad de las decisiones tenemos

$$g_t(\hat{x}_t) = B_t(\hat{x}_t, \hat{u}_t, \cdots, \hat{u}_{T-1}) = V_t(\hat{x}_t).$$

A partir de la hipótesis $(H)_A$ deducimos que las funciones g_t son afínes. Por lo tanto para todo $x \in X_t$ se cumple

$$g_t(x) - g_t(\hat{x}_t) = \langle \nabla g_t(\hat{x}_t); x - \hat{x}_t \rangle.$$

Con este resultado y la función g_t en la definición de $\rho_t(\hat{x}_t)$ tenemos

$$\begin{array}{rcl} \rho_t(\hat{x}_t) & = & \sup_{u_t \in U_t} [f_t^0(\hat{x}_t, u_t) + B_T^{t+1}(f_t(\hat{x}_t, u_t); \hat{u}_{t+1}, \cdots, \hat{u}_{T-1})] \\ & = & \sup_{u_t \in U_t} [f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t))] \\ & = & \sup_{u_t \in U_t} [f_t^0(\hat{x}_t, u_t) + \langle \nabla g_{t+1}(\hat{x}_{t+1}); f_t(\hat{x}_t, u_t) - \hat{x}_{t+1} \rangle + g_{t+1}(\hat{x}_{t+1})] \\ & = & C(\hat{x}_t) + \sup_{u_t \in U_t} [f_t^0(\hat{x}_t, u_t) + \langle \nabla g_{t+1}(\hat{x}_{t+1}); f_t(\hat{x}_t, u_t) \rangle], \end{array}$$

donde $C(\hat{x}_{t+1}) = g_{t+1}(\hat{x}_{t+1}) - \langle \nabla g_{t+1}(\hat{x}_{t+1})\ ;\ \hat{x}_{t+1}\rangle$.

Definiendo $\hat{\psi}_t := \nabla g_{t+1}(\hat{x}_{t+1})$, y teniendo en cuenta la definición del Hamiltoniano.

$$\begin{array}{rcl} \rho_t(\hat{x}_t) & = & C(\hat{x}_t) + \sup_{u_t\in U_t}[f_t^0(\hat{x}_t,u_t) + \langle \hat{\psi}_t; f_t(\hat{x}_t,u_t)\rangle] \\ & = & C(\hat{x}_t) + \sup_{u_t\in U_t} H_t(\hat{x}_t,u_t,\hat{\psi}_t) \end{array}$$

Teniendo en cuenta que $\rho_t(\hat{x}_t) = V_t(\hat{x}_t)$, deducímos

$$\begin{array}{rcl} V_t(\hat{x}_t) - C(\hat{x}_t) & = & \rho_t(\hat{x}_t) - C(\hat{x}_t) \\ & = & \sup_{u_t\in U_t} H_t(\hat{x}_t,u_t,\hat{\psi}_t). \end{array}$$

Calculando $V_t(\hat{x}_t) - C(\hat{x}_t)$ tenemos

$$\begin{array}{rcl} V_t(\hat{x}_t) - C(\hat{x}_t) & = & g_t(\hat{x}_t) - g_{t+1}(\hat{x}_{t+1}) + \langle \nabla g_{t+1}(\hat{x}_{t+1}); \hat{x}_{t+1}\rangle) \\ & = & f_t^0(\hat{x}_t,\hat{u}_t) + \langle \hat{\psi}_t, \hat{x}_t\rangle) \\ & = & H_t(\hat{x}_t,\hat{u}_t,\hat{\psi}_t). \end{array}$$

Por lo tanto

$$H_t(\hat{x}_t,\hat{u}_t,\hat{\psi}_t) = \sup_{u\in U_t} H_t(\hat{x},u,\hat{\psi}_t).$$

Así pues $\hat{u}_t \in U_t$ es una solución óptima del problema $(PM)_t$.

Finalmente por la la definición realizada $\hat{\psi}_t := \nabla g_{t+1}(\hat{x}_{t+1})$ y la Proposición 4.1.1 tales variables son solución del sistema (EA). ■

Observación. El resultado anterior es válido tanto si $\hat{x}_t \in \operatorname{int} X_t$, como si esta en la frontera ($\hat{x}_t \in \partial X_t$). Tampoco depende de la forma de los conjuntos U_t, estos pueden ser discretos, cerrados, abiertos, etc. Estos hechos confieren a este resultado una generalidad sorprendente.

Habiendo probado que el PFP es una condición necesaria de optimalidad para problemas de este tipo. Nos preguntamos si será también una condición suficiente. La respuesta es no. Veamos a continuación un contraejemplo.

Contraejemplo: Condición Suficiente. Para el problema $\mathcal{P}_T(\hat{x}_0)$, consideremos $T = 2$, $\hat{x}_0 := 1 \in X_0 := \mathbb{R}$, $X_1 = X_2 := \mathbb{R}$ y $U_1 = U_2 := \{1,2\}$. Definamos las funciones de evolución

$$\begin{array}{rccl} f_0: & X_0\times U_0 & \to & \mathbb{R} \\ & (x,u) & \mapsto & x\cdot u, \end{array} \qquad y \qquad \begin{array}{rccl} f_1: & X_1\times U_1 & \to & \mathbb{R} \\ & (x,u) & \mapsto & x\cdot u, \end{array}$$

y las funciones de beneficio $f_0^0 : X_0\times U_0 \to \mathbb{R}$ y $f_1^0 : X_1\times U_1 \to \mathbb{R}$.

$$f_0^0(x,u) := \left(\frac{3}{2} - \frac{u}{2}\right)x, \qquad y \qquad f_1^0(x,u) := [2(u-1)]x + [2 - 3(u-1)].$$

La función $f_2^0 \equiv 0$. Es claro que el problema satisface las hipótesis $(H)_A$ y $(Inclu)$.

Puesto que sólo se usan 2 controles, no es dificil calcular manualmente todas las posibilidades (apenas 4).

Encontramos que **la solución óptima es** $(\hat{u}_0 = 2, \hat{u}_1 = 2)$, con su proceso asociado $(\hat{x}_0 = 1, \hat{x}_1 = 2, \hat{x}_2 = 4)$. Dando un beneficio total de 3.5.

De otro lado **la sucesión de decisiones** $(\bar{u}_0 = 1, \bar{u}_1 = 1)$, con su proceso asociado $(\hat{x}_0 = 1, \bar{x}_1 = 1, \bar{x}_2 = 1)$, **no es solución** óptima. Sin embargo veremos que **satisface el PFP**.

Primero calculemos las variables adjuntas

$$\begin{aligned} \bar{\psi}_1 &= \tfrac{d}{dx} f_2^0(\bar{x}_2) & &= \tfrac{d}{dx} f_2^0(1) = 0, \\ \\ \bar{\psi}_0 &= \tfrac{d}{dx} H_1(\bar{x}_1, \bar{u}_1, \bar{\psi}_1) \\ &= \tfrac{d}{dx} f_1^0(\bar{x}_1, \bar{u}_1) & &= 2(\bar{u}_1 - 1) = 0. \end{aligned}$$

Con estas variables adjuntas calculamos los hamiltonianos,

$$H_1(\bar{x}_1, u; \bar{\psi}_1) = f_1^0(\bar{x}_1, u) \qquad \text{y} \qquad H_0(\hat{x}_0, u, \bar{\psi}_0) = f_0^0(\hat{x}_0, u).$$

Veamos que las decisiones $\bar{u}_0$ y $\bar{u}_1$ son soluciones de los subproblemas $(PM)_t$. En primer lugar tenemos

$$\begin{aligned} f_1^0(\bar{x}_1, 1) = 2 &\geq f_1^0(\bar{x}_1, 2) = 1, \\ H_1(\bar{x}_1, 1, \bar{\psi}_1) &\geq H_1(\bar{x}_1, 2, \bar{\psi}_1). \end{aligned}$$

por lo cual, como $U_1 = \{1, 2\}$, se tiene que $\bar{u}_1 = 1$ es solución de (PM_1)

$$\max_{u \in U_1} H_1(\bar{x}_1, u, \bar{\psi}_1) \tag{$(PM)_1$}$$

Similarmente desde que,

$$\begin{aligned} f_1^0(\bar{x}_1, 1) = 1 &\geq f_1^0(\bar{x}_1, 2) = 1/2, \\ H_0(\hat{x}_0, 1, \bar{\psi}_0) &\geq H_1(\hat{x}_0, 2, \bar{\psi}_0), \end{aligned}$$

la decisión $\bar{u}_0 = 1$ es solución del subproblema $(PM)_0$

De estos cálculos **la sucesión de decisiones** $\bar{u}_0 = 1, \bar{u}_1 = 1$ **satisface la condición necesaria PFP pero no es una solución óptima.** Por lo que el PFP no es una condición suficiente para los problemas tratados en esta sección (afines en la variable de estado).

Finalmente, para resaltar con este ejemplo que el PFP sí es una condición necesaria, observemos que la solución óptima $(\hat{u}_0 = 2, \hat{u}_1 = 2)$, si satisface la condición necesaria. Calculando las variables adjuntas se obtiene $\hat{\psi}_1 = 0$ y $\hat{\psi}_0 = 2$, con las cuales se obtiene que

$$H_0(1, 2, \hat{\psi}_0) = \tfrac{1}{2} + 4 \quad \geq \quad H_0(1, 1, \hat{\psi}_0) = 1 + 2,$$

$$H_1(2,2,\hat{\psi}_1) = 3 \quad \geq \quad H_0(1,1,\hat{\psi}_0) = 2.$$

Vemos pues que el Principio Fuerte de Pontryaguin no es una condición suficiente en este caso.

La pregunta natural ahora es saber en que condiciones el PFP es también una condición suficiente. Al intentar realizar investigaciones destinadas a resolver esta pregunta encontramos que existe una relación profunda entre el PFP y la Programación Dinámica. Estos resultados los exponemos en la siguiente subsección.

4.1.3 Programación Dinámica y el PFP

Comenzemos nuestro estudio con la siguiente definición.

Definición 4.1.1 (Sucesión PFP) *Para el problema $\mathcal{P}_T(\hat{x}_0)$, diremos que una sucesión de decisiones $(\hat{u}_0,\cdots,\hat{u}_{T-1}) \in \mathrm{Adm}_{\mathrm{T}}(\hat{\mathrm{x}}_0)$ con su proceso asociado $(\hat{x}_0,\cdots,\hat{x}_T)$ es una* **sucesión PFP**, *si existen $\hat{\psi}_0,\cdots,\hat{\psi}_{T-1} \in \mathbb{R}^n$ tales que:*

1. *Para $t = 1,\cdots,N-1$,*

$$\hat{\psi}_{t-1} = \nabla_x(H_t(\hat{x}_t,\hat{u}_t,\hat{\psi}_t), \qquad (EA)$$

con condición final, $\hat{\psi}_{N-1} = \nabla_x f_T^0(\hat{x}_T)$.

2. *Para $t = 0,\cdots,T-1$, $\hat{u}_t$ es* **solución** *del subproblema*

$$\max_{u\in U_t} H_t(\hat{x},u,\hat{\psi}_t), \qquad (PM)_t$$

donde $H_t(x_t,u_t,\psi_t) = f_t^0(x_t,u_t) + \langle \psi_t, f_t(x_t,u_t)\rangle$ es el hamiltoniano.

Proposición 4.1.3 (Caracterización de sucesiones PFP) *Sea el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo las hipótesis del Teorema 4.1.1, una sucesión de decisiones $(\hat{u}_0,\cdots,\hat{u}_{T-1})$ será* **PFP** *si y sólo si para $t = 0,\cdots,T-1$, $\hat{u}_t$ es* **solución óptima** *del subproblema*

$$\max_{u_t\in U_t} B_T^t(\hat{x}_t;u_t,\hat{u}_{t+1},\cdots,\hat{u}_{T-1}),$$

lo que es equivalente a que para $t = 0,\cdots,T-1$,

$$\rho_t(\hat{x}_t) = B_T^t(\hat{x}_t;\hat{u}_t,\cdots,\hat{u}_{T-1}),$$

donde ρ_t son las funciones asociadas a $(\hat{u}_0,\cdots,\hat{u}_{T-1})$, y $(\hat{x}_0,\cdots,\hat{x}_T)$ es su proceso asociado.

Prueba. Consideremos para $t = 0, \cdots, T-1$, las funciones g_t asociadas a la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, tenemos

$$\begin{aligned} B_T^t(\hat{x}_t; u_t, \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) &= f_t^0(\hat{x}_t, u_t) + B_T^{t+1}(f_t(\hat{x}_t, u_t); \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) \\ &= f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t)). \end{aligned}$$

Por la hipótesis $(H)_A$, las funciones g_t son afines y por ello para todo $x \in X_t$ se cumple

$$g_t(x) - g_t(\hat{x}_t) = \langle \nabla g_t(\hat{x}_t); x_t - \hat{x}_t \rangle.$$

Considerando esto y la definición del Hamiltoniano obtenemos

$$\begin{aligned} B_T^t(\hat{x}_t; u_t, \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) &= f_t^0(\hat{x}_t, u_t) + \langle \nabla g_{t+1}(\hat{x}_{t+1}); f_t(\hat{x}_t, u_t) \rangle + C(\hat{x}_{t+1}) \\ &= H_t(\hat{x}_t, u_t, \nabla g_{t+1}(\hat{x}_{t+1})) + C(\hat{x}_{t+1}) \\ &= H_t(\hat{x}_t, u_t, \hat{\psi}_t) + C(\hat{x}_{t+1}), \end{aligned}$$

donde $\hat{\psi}_t = \nabla g_{t+1}(\hat{x}_{t+1})$ y $C(\hat{x}_{t+1}) = g_{t+1}(\hat{x}_{t+1}) - \langle \nabla g_{t+1}(\hat{x}_{t+1}); \hat{x}_{t+1} \rangle)$.

Así pues, salvo la constante $C(\hat{x}_{t+1})$, tenemos la siguiente equivalencia

$$\max_{u \in U_t} H_t(\hat{x}, u, \hat{\psi}_t) \qquad \equiv \qquad \max_{u_t \in U_t} B_T^t(\hat{x}_t; u_t, \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}).$$

De la definición de las funciones ρ_t tenemos las otras conclusiones de la proposición. ■

Interpretando el resultado anterior adecuadamente encontramos una relación con la Programación Dinámica.

En efecto del Corolario 2.2.1 una sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ es óptima si y sólo si para $t = 0, \cdots, T-1$, cada $\hat{u}_t$ es solución del problema

$$\max_{u_t \in U_t} f_t^0(\hat{x}_t, u_t) + V_{t+1}(f_t(\hat{x}_t, u_t)).$$

De otro lado para $t = 0, \cdots, T-1$,

$$\begin{aligned} B_T^t(\hat{x}_t; u_t, \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) &= f_t^0(\hat{x}_t, u_t) + B_T^t(f_t(\hat{x}_t, u_t); \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) \\ &= f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t)). \end{aligned}$$

Encontramos claramente una similitud entre los problemas

$$\max_{u_t \in U_t} [f_t^0(\hat{x}_t, u_t) + V_{t+1}(f_t(\hat{x}_t, u_t))] \qquad \sim \qquad \max_{u_t \in U_t} [f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t))].$$

Desde la generalidad de la función valor se ve claramente que el segundo problema es un caso particular del primero, es decir, si $\hat{u}_t$ es solución del primero entonces lo será del segundo.

De la equivalencia anterior si queremos que el principio fuerte (PFP) sea una condición de optimalidad suficiente debemos exigir condiciones que aseguren que a partir del segundo problema se deduzca una solución del primero, pues las soluciones del primero corresponden a óptimos del problema $P_T(x_0)$ (Ecuaciones de Bellman). En esta dirección tenemos la siguiente proposición.

Proposición 4.1.4 *Para el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo las hipótesis del Teorema 4.4.1, una sucesión* **PFP** $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ *será* **solución óptima** *si y sólo si para cada $t = 0, \cdots, T-1$, se cumple:*
Para todo $u_t \in U_t$,

$$V_{t+1}(f_t(\hat{x}_t, u_t)) - g_{t+1}(f_t(\hat{x}_t, u_t)) \leq [B_T^t(\hat{x}_t; \hat{u}_t, [\hat{u}_j]_j)] - [B_T^t(\hat{x}_t; u_t, [\hat{u}_j]_j)],$$

donde $[\hat{u}_j]_j := (\hat{u}_{t+1}, \cdots, \hat{u}_{T-1})$.

Prueba. Desde la definición de las funciones g_t,

$$B_T^t(\hat{x}_t; u_t, [\hat{u}_j]_j) = f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t)).$$

Con esta observación y haciendo algunos cálculos, obtenemos que la desigualdad de la proposición es equivalente a

$$V_{t+1}(f_t(\hat{x}_t, u_t)) + f_t^0(\hat{x}_t, u_t) \leq [f_t^0(\hat{x}_t, \hat{u}_t) + g_{t+1}(f_t(\hat{x}_t, \hat{u}_t))] = B_T^t(\hat{x}_t; \hat{u}_t, \cdots, \hat{u}_{T-1}).$$

Así pues, la condición será equivalente a:
Para todo $u_t \in U_t$ se cumple

$$V_{t+1}(f_t(\hat{x}_t, u_t)) + f_t^0(\hat{x}_t, u_t) \leq B_T^t(\hat{x}_t; \hat{u}_t, \cdots, \hat{u}_{T-1}).$$

Usando la definición de la función valor V_t y las Ecuaciones de Bellman, lo anterior es equivalente a

$$\begin{array}{rclcl} V_t(\hat{x}_t) & = & \sup_{u_t \in U_t}(V_{t+1}(f_t(\hat{x}_t, u_t)) + f_t^0(\hat{x}_t, u_t)) & \leq & \\ & \leq & B_T^t(\hat{x}_t; \hat{u}_t, \cdots, \hat{u}_{T-1}) & \leq & V_t(\hat{x}_t). \end{array}$$

Por lo tanto la condición exigida en la proposición es equivalente a pedir que para $t = 0, \cdots, T-1$ se cumpla

$$V_t(\hat{x}_t) = B_T^t(\hat{x}_t; \hat{u}_t, \cdots, \hat{u}_{T-1}).$$

Del Corolario 2.2.1, esto es equivalente a que $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ sea una solución óptima del problema $\mathcal{P}_T(\hat{x}_0)$. ■

La proposición anterior nos sirve para observar la relación en la que debe estar la función g_t respecto a la función valor. Se deduce que si esta suficientemente "cerca" entonces podremos asegurar que la sucesión PFP es una solución del problema $\mathcal{P}_T(\hat{x}_0)$.

Como corolario de la proposición anterior mostraremos que para cierto tipo de problemas la condición PFP es también suficiente.

Consideremos problemas $\mathcal{P}_T(\hat{x}_0)$ que satisfagan la siguiente **hipótesis**:

"Exista una solución $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ *que satisface: Para todo* $x_t \in X_t$ *la sucesión* $(\hat{u}_t, \cdots, \hat{u}_{T-1})$ *es la única solución óptima del subproblema* $\mathcal{P}_T^t(x_t)$ *".*

Por ello la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ puede interpretarse como desiciones óptimas independientes del estado x_t al que se aplican, sólo dependen del momento en el que se aplican.

Para problemas de este tipo se tiene la siguiente equivalencia entre las ecuaciones de Bellman y las del PFP, en particular en este caso, PFP será una condición suficiente.

Proposición 4.1.5 (Equivalencia: PFP y las Ecuaciones de Bellman) *Para el problema* $\mathcal{P}_T(\hat{x}_0)$ *satisfaciendo las hipótesis del Teorema 4.1.1 y adicionalmente la hipótesis anterior. Sea una sucesión de decisiones* $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ *con su proceso asociado* $(\hat{x}_0, \cdots \hat{x}_T)$*, y sus variables adjuntas* $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1}$*. Se tiene la siguiente equivalencia*
Para todo $t = 0, \cdots, T-1$, $\hat{u}_t$ *es* ***solución*** *de*

$$\max_{u \in U_t} H_t(\hat{x}, u, \hat{\psi}_t), \tag{PM}_t$$

si y sólo si es solución de

$$V_t(\hat{x}_t) = \max_{u_t \in U_t} [f_t^0(\hat{x}_t, u_t) + V_{t+1}(f_t(\hat{x}_t, u_t))]. \tag{EB}_t$$

Por lo tanto $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ *es una sucesión PFP si y sólo si es una solución óptima de* $\mathcal{P}_T(\hat{x}_0)$*.*

Prueba. Sabemos que las Ecuaciones de Bellman caracterizan las soluciones del problema. Por hipótesis estamos suponiendo que existe una única solución, denotemósla $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, más aún sus subsucesiones son soluciones únicas de todos los subproblemas $P_T^t(x_t)$ para $x_t \in X_t$ fijo y arbitrario, por ello

$$V_t(x_t) = B_T^t(x_t, \hat{u}_t, \cdots, \hat{u}_{T-1}).$$

Probemos la primera implicación.

($\Rightarrow$) Sea $(\bar{u}_0, \cdots, \bar{u}_{T-1})$ una sucesión PFP, entonces para $t = 0, \cdots, T-1$, $\bar{u}_t$ es solución del problema

$$\max_{u_t \in U_t} H_t(\bar{x}_t, u_t; \bar{\psi}_t).$$

De la Proposición 4.1.3 esto es equivalente a que $\bar{u}_t$ es solución del problema

$$\max_{u_t \in U_t} B_T^t(\bar{x}_t, u_t, \bar{u}_{t+1}, \cdots, \bar{u}_{T-1}).$$

Por lo tanto para todo $u_t \in U_t$

$$0 \leq B_T^t(\bar{x}_t; \bar{u}_t, \bar{u}_{t+1}, \cdots, \bar{u}_{T-1}) - B_T^t(\bar{x}_t; u_t, \bar{u}_{t+1}, \cdots, \bar{u}_{T-1})$$

Con todo esto, para probar la implicación sólo nos faltaría probar que para todo $u_t \in U_t$ se cumple

$$0 = V_{t+1}(f_t(\hat{x}_t, u)) - g_{t+1}(f_t(\hat{x}_t, u)),$$

donde las funciones g_t son las funciones asociadas a la sucesión $(\bar{u}_0, \cdots, \bar{u}_{T-1})$.

Probémoslo por inducción: Dado $u \in U_{T-1}$ fijo y arbitrario, tenemos por sus definiciones

$$V_T(f_{T-1}(\bar{x}_{T-1}, u)) = f_T^0(f_{T-1}(\bar{x}_{T-1}, u)) = g_T(f_{T-1}(\bar{x}_{T-1}, u)).$$

Supongamos que hayamos probado hasta "r" es decir para toda $u_r \in U_r$

$$V_{r+1}(f_r(\bar{x}_r, u)) = g_{r+1}(f_r(\bar{x}_r, u)).$$

Entonces tendremos

$$\begin{aligned} f_r^0(\bar{x}_r, u) + V_{r+1}(f_r(\bar{x}_r, u)) &= f_r^0(\bar{x}_r, u) + g_{r+1}(f_r(\bar{x}_r, u)) \\ &= B_T^r(\bar{x}_r, u_r, \bar{u}_{r+1}, \cdots, \bar{u}_{T-1}). \end{aligned}$$

Por lo tanto son equivalentes

$$\max_{u \in U_t}[f_r^0(\bar{x}_r, u) + V_{r+1}(f_r(\bar{x}_r, u))] \qquad \equiv \qquad \max_{u \in U_t} B_T^r(\bar{x}_r, u_r, \bar{u}_{r+1}, \cdots, \bar{u}_{T-1}).$$

Por ello $\bar{u}_r$ es solución del primer problema. De las Ecuaciones de Bellman, deducímos

$$V_r(\bar{x}_r) = \max_{u \in U_t}[f_r^0(\bar{x}_r, u) + V_{r+1}(f_r(\bar{x}_r, u))] = B_T^r(\bar{x}_r, \bar{u}_r, \bar{u}_{r+1}, \cdots, \bar{u}_{T-1}).$$

Así pues $(\bar{u}_r, \bar{u}_{r+1}, \cdots, \bar{u}_{T-1})$ es solución del problema $\mathcal{P}_T(\bar{x}_r)$.

Por hipótesis sólo existe una única solución que es $(\hat{u}_r, \hat{u}_{r+1}, \cdots, \hat{u}_{T-1})$. Por lo tanto $\bar{u}_s = \hat{u}_s$, con $s = r, \cdots, T-1$.

Ahora estudiemos el problema para "$r-1$". Tenemos que probar que para toda $u \in U_{r-1}$,

$$V_r(f_{r-1}(\bar{x}_{r-1}, u)) = g_r(f_{r-1}(\bar{x}_{r-1}, u)).$$

Dada $u_{r-1} \in U_{r-1}$ fijo y arbitraria, $(\hat{u}_r, \hat{u}_{r+1}, \cdots, \hat{u}_{T-1})$, es la única solución de $\mathcal{P}_T(f_{r-1}(\bar{x}_{r-1}, u))$. Por lo tanto deducímos

$$\begin{aligned} V_r(f_{r-1}(\bar{x}_{r-1}, u)) &= B_T^r(f_{r-1}(\bar{x}_{r-1}, u); \hat{u}_r, \hat{u}_{r+1}, \cdots, \hat{u}_{T-1}) \\ &= B_T^r(f_{r-1}(\bar{x}_{r-1}, u); \bar{u}_r, \bar{u}_{r+1}, \cdots, \bar{u}_{T-1}) \\ &= g_r(f_{r-1}(\bar{x}_{r-1}, u)). \end{aligned}$$

Como $u_r \in U_r$ es arbitraria, deducimos que para toda $u_r \in U_r$:

$$V_r(f_{r-1}(\bar{x}_{r-1}, u)) = g_r(f_{r-1}(\bar{x}_{r-1}, u)).$$

Continuando de esta forma podemos deducir que $\bar{u}_s = \hat{u}_s$, para $s = 0, \cdots, u_{T-1}$. Y por lo tanto la sucesión PFP es la solución óptima.

($\Leftarrow$)

Es claro que si $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ es solución de las Ecuaciones de Bellman entonces es solución del problema, por lo tanto, del Teorema 4.1.1 satisface la condición necesaria PFP. En consecuencia para todo $t = 0, \cdots, T-1$, $\hat{u}_t$ es solución de

$$\max_{u \in U_t} H_t(\hat{x}, u, \hat{\psi}_t). \qquad (PM)_t$$

■

Observación. Con el resultado anterior encontramos que existen casos en los cuales el PFP es una condición suficiente. ¿Habrá otros casos en los que también lo sea? En esta línea nos preguntamos si con condiciones de concavidad respecto a las variables de estado "x" y control "u", habra suficiencia. La respuesta es no. En el contraejemplo para la suficiencia de PFP, las funciones f_t^0 y f_t lineales para (x, u), sin embargo no hay suficiencia.

Estudiemos dos ejemplos en los que aplicaremos los resultados obtenidos en esta última parte.

Ejemplo 1: Decisiones Independientes. Consideremos un problema $\mathcal{P}_T(\hat{x}_0)$ tal que los beneficios de una decisión no dependen del estado x_t, es decir $f_t^0(x_t, u_t) = f_t^0(u_t)$ y la función $f_T^0 \equiv 0$.

Llamaremos a este problema de **decisiones independientes del estado** porque en cada momento decidimos sin importar el estado en el que se encuentra x_t para tomar la decisión óptima.

Dada una sucesión óptima $(\hat{u}_0, \cdots, \hat{u}_{T-1})$. al resolver el sistema adjunto (EA). Encontramos que $\hat{\psi}_0 = \cdots = \hat{\psi}_{T-1} = 0$.

Desde el PFP tenemos que para $t = 0, \cdots, T-1$, $\hat{u}_t$ es solución del problema

$$\max_{u \in U_t} H_t(\hat{x}, u, \hat{\psi}_t) = f_t^0(u_t) + \langle 0, f_t(\hat{x}_t, u_t) \rangle = \max_{u \in U_t} f_t^0(u_t).$$

Este resultado era intuitivo. En efecto si los beneficios de las decisiones no dependen del estado x_t en el que se toman, entonces lo mejor en cada instánte es tomar aquella decisión que nos de mayor beneficio. Este tipo de problemas nos hace recordar la vieja máxima: **"Haz lo mejor cada momento y alcanzarás la victoria"**. Sin embargo, pese a la máxima, hemos de decir que en general tal recomendación no es válida. Pues para problemas en los que el beneficio si depende del estado "x_t", lo mejor no es optimizar el beneficio instantáneo. Sino tomar decisiones que nos lleven continuamente a mejores "escenarios futuros". Más precisamente la decisión debe atender dos objetivos: 1) de un lado maximizar el beneficio actual $f_t^0(x_t, u_t)$ y, 2) el beneficio potencial máximo esperado para el estado que se alcanza luego de la decisión, $V_{t+1}(x_{t+1})$. Esta conclusión nace de analizar las Ecuaciones de Bellman.

Ejemplo 2: Problema Lineal. Consideremos el problema $\mathcal{P}_T(\hat{x}_0)$ para el cual: $T \in \mathbb{N}$, $x_0 = \bar{x}_0 \in \mathbb{R}^n$, $x_t \in X_t := \mathbb{R}^n, u_t \in U_t := U = B_1[0] \subseteq \mathbb{R}^m$

Para $t = 0, \cdots, T-1$, tenemos las funciones

$$\begin{array}{rccc} f_t: & \mathbb{R}^n \times B_1[0] & \rightarrow & \mathbb{R}^n \\ & (x,u) & \mapsto & Ax + Bu \end{array} \qquad \begin{array}{rccc} f_t^0: & \mathbb{R}^n \times B_1[0] & \rightarrow & \mathbb{R} \\ & (x,u) & \mapsto & \langle a, x\rangle + \langle b, u\rangle, \end{array}$$

donde $a \in \mathbb{R}^n, b \in \mathbb{R}^m$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ y la función $f_T^0 : \mathbb{R}^n \to \mathbb{R}$, será $f_T^0 \equiv 0$.

A problemas de este tipo les denominamos **problemas lineales** .

Para hallar la solución usando la Programación Dinámica hemos de pedir que se satisfaga las siguientes condiciones adicionales:

1. $\|b\| \neq 0, \quad b + B^t a \neq 0.$

2. Para $k = 1, \cdots, T-2$,

$$b + B^t(I + A^t + \ldots + (A^t)^{T-k-2})a \neq 0.$$

Usemos en primer lugar la condición necesaria PFP, para hallar los candidatos a solución óptima. Veremos que esta metodología es mucho más eficiente que la Programación Dinámica.

Primero calculemos las variables adjuntas. Veremos que son las mismas para toda sucesión de decisiones $(\hat{u}_0, \cdots, \hat{u}_{T-1})$. Tomemos una de ellas, con su proceso asociado $(\hat{x}_0, \cdots, \hat{x}_T)$. Primero es claro $\psi_{T-1} = \frac{d}{dx} f_T^0 = 0$,

$$\psi_{T-2} = \nabla_x H_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1}, 0) = \nabla_x(\langle a, x\rangle + \langle b, u\rangle) = a,$$

$$\psi_{T-3} = \nabla_x H_{T-1}(\hat{x}_{T-1}, \hat{u}_{T-1}, a) = \nabla_x(\langle a, x\rangle + \langle b, u\rangle + \langle a, Ax + Bu\rangle) = a + a \circ A.$$

Inductivamente tenemos

$$\psi_{T-k} = a + a \circ A + a \circ A^2 + \cdots + a \circ A^{k-2}.$$

Vemos pues que las soluciones sólo dependen del vector a y de la matriz A.

Sea $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ una sucesión PFP. Aplicando el principio del máximo para $r = 0, \cdots, T-2$, $\hat{u}_r$ es solución de

$$\max_{u_r \in U_r} H_r(\hat{x}_r, u_r, \psi_r).$$

Con las variables adjuntas calculadas, son equivalentes:

$$\max_{u_r \in U_r} \{\langle a, \hat{x}_r\rangle + \langle b, \hat{u}_r\rangle + +\langle (a + a \circ A + \cdots + a \circ A^{T-r-2}); Ax_r + Bu_r\rangle\},$$

$$K(\hat{x}_r) + \max_{u_r \in U_r} \{\langle b, u_r\rangle + \langle (a + a \circ A + \cdots + a \circ A^{T-r-2}); Bu_r\rangle\},$$

$$K(\hat{x}_r) + \max_{u_r \in U_r} \{\langle b + B^t(a + a \circ A + \cdots + a \circ A^{T-r-2}); u_r\rangle\},$$

donde $K(\hat{x}_r) := \langle a, \hat{x}_r\rangle + \langle b + B^t(a + a \circ A + \cdots + a \circ A^{T-r-2}); Ax_r\rangle$, y B^t es la matriz traspuesta de B.

Como $U_r = B_1[0]$ es claro que la única solución del problema anterior es el vector

$$\hat{u}_r = \frac{b + B^t(a + a \circ A + \cdots + a \circ A^{T-r-2})}{\|b + B^t(a + a \circ A + \cdots + a \circ A^{T-r-2})\|}.$$

Desde que $\psi_{T-1} = 0$ de modo similar se calcula $\hat{u}_{T-1} = \frac{b}{\|b\|}$.

Notemos que la hipótesis inicial permite asegurar que estos vectores están bien definidos (su norma es distinta de cero).

Así pues tenemos una única sucesión PFP. Por otra parte del Teorema 1.2.2, nos asegura que existe solución óptima para este problema. Por lo tanto de la condición necesaria PFP afín (Teorema 4.1.1) esta solución es una sucesión PFP. Por lo tanto nuestra sucesión PFP es solución óptima.

Por otra parte T es arbitrario, y también lo es el estado de inicio $\hat{x}_0$, se sigue que la sucesión $\hat{u}_r, \cdots, \hat{u}_{T-1}$, es solución óptima del problema $\mathcal{P}_T(x_r)$ para todo $x_r \in X_r$, más áun es única pues nuevamente es la única sucesión PFP. Vemos entonces que este problema cumple las condiciones exigidas para la Proposición 4.1.5.

Por otra parte podemos usar también la Programación Dinámica para hallar las soluciones de este problema. El resultado es el mismo. Existe una única solución que satisface las ecuaciones de Bellman y coincide con la que hemos hallado.

Al hacer los cálculos la Programación Dinámica exigió más trabajo que el PFP. Esto se debe a que en el PFP se calcula directamente las variables adjuntas y luego hallar las soluciones óptimas de cada problema $(PM)_t$ es inmediato. Mientras que en la Programación Dinámica se exige hallar funciones recursivamente y requiere ir resolviendo una tras otra para conseguir la solución final.

4.1.4 Generalización del PFP Afín

En esta subsección las hipótesis de afinidad respecto a la variable de estado "x" ya no se mantendrán, sólo permanecerán las hipótesis (Inclu) y la diferenciabilidad respecto a la variable "x" para las funciones f_t y f_T^0.

Buscaremos obtener una generalización de la metología utilizada en la primera parte de esta sección con la intención de estudiar problemas más generales.

A parte de las hipótesis de afinidad dos fueron los pilares sobre los cuales conseguimos obtener el Teorema PFP-Afin (Teorema 4.1.1).

1. La caracterización de las variables adjuntas, Proposición 4.1.1,

2. La relación entre el PFP, y los problemas siguientes (Proposición 4.1.3):

$$\max_{u_t \in U_t} B_T^t(\hat{x}_t; u_t, \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}).$$

Ambas conclusiones no dependen de la hipótesis de afinidad $(H)_A$. Por lo tanto siguen siendo ciertas en el caso más general de este subsección. De estas observaciones tenemos los siguientes lemas.

Lema 4.1.2 *Para el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo la hipótesis (inclu), sea la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ con su proceso asociado $(\hat{x}_0, \cdots, \hat{x}_T)$, una solución óptima. Entonces se cumple:*
Para $r = 0, \cdots, T-1$, $\hat{u}_r$ es solución del problema

$$\max_{u_r \in U_r} [f_t^0(\hat{x}_r, u_r) + g_{r+1}(f_r(\hat{x}_r, u_r))],$$

donde g_r son las funciones asociadas $(\hat{u}_0, \cdots, \hat{u}_{T-1})$.

Prueba. Tomemos $r = 0, \cdots, T-1$, fijo y arbitrario. Desde la optimalidad de la sucesión, las funciones valor y la definición de la función g_{r+1} tenemos

$$f_t^0(\hat{x}_r, \hat{u}_r) + g_{r+1}(f_r(\hat{x}_r, \hat{u}_r)) = V_r(\hat{x}_r) = \max_{u_r \in U_r} [f_t^0(\hat{x}_r, u_r) + V_{r+1}(f_r(\hat{x}_r, u_r))].$$

Desde que $g_{r+1}(f_r(\hat{x}_r, u_r)) \leq V_{r+1}(f_r(\hat{x}_r, u_r))$, se deduce el lema. ∎

Lema 4.1.3 *Consideremos el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo las hipótesis de la Proposición 4.1.1. Para $t = 0, \cdots, T-1$, se tiene la equivalencia*

$$\max_{u_t \in U_t} \{f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t))\} \quad \equiv \quad \max_{u_t \in U_t} \{H_t(\hat{x}_t, u_t, \hat{\psi}_t) + o_t(u_t)\},$$

donde la función $o_t : U_t \to \mathbb{R}$ es definida según,

$$o_t(u_t) := g_{t+1}(f_t(\hat{x}_t, u_t)) - g_{t+1}(f_t(\hat{x}_t, \hat{u}_t)) - \langle \hat{\psi}_t, (f_t(\hat{x}_t, u_t) - f_t(\hat{x}_t, \hat{u}_t)) \rangle,$$

con $\hat{\psi}_t := \nabla_x g_{t+1}(\hat{x}_{t+1})$, y $H_t(x_t, u_t; \psi_t) := f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$ el hamiltoniano.

Prueba. Usando la definición de la función $o(t)$ y realizando algunos cálculos se obtiene facilmente que $H_t(\hat{x}_t, u_t, \hat{\psi}_t) + o_t(u_t)$, es igual a

$$f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t)) - C(\hat{x}_t),$$

donde $C(\hat{x}_t) = g_{t+1}(\hat{x}_{t+1}) - \langle\hat{\psi}_t, f_t(\hat{x}_t, \hat{u}_t)\rangle$. Entonces ambos términos son iguales salvo una constante que depende de $\hat{x}_t$ y por ello los problemas son equivalentes. ■

A partir de estos dos últimos lemas podemos demostrar que el PFP es válido cuando se satisface una condición adicional.

Proposición 4.1.6 *Consideremos el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo las hipótesis de la Proposición 4.1.1. Sea $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ una solución óptima, con su proceso asociado $(\hat{x}_0, \cdots, \hat{x}_T)$, tal que satisface la siguiente condición:*
Para toda $u_t \in U_t$,

$$o_t(u_t) \geq 0,$$

donde $o_t : U_t \to \mathbb{R}$ es definida según

$$o_t(u_t) := g_{t+1}(f_t(\hat{x}_t, u_t)) - g_{t+1}(f_t(\hat{x}_t, \hat{u}_t)) - \langle\hat{\psi}_t, [f_t(\hat{x}_t, u_t) - f_t(\hat{x}_t, \hat{u}_t)]\rangle,$$

con $\hat{\psi}_t := \nabla_x g_{t+1}(\hat{x}_{t+1})$.

Entonces $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ satisface el PFP (es una sucesión PFP).

Prueba. En efecto de los Lemas 4.1.2 y 4.1.3, para $t = 0, \cdots, T-1$, $\hat{u}_t$ es solución del subproblema

$$\max_{u_t \in U_t} [f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t))] \quad \equiv \quad \max_{u_t \in U_t} [H_t(\hat{x}_t, u_t, \hat{\psi}_t) + o_t(u_t)].$$

Por lo tanto para toda $u_t \in U_t$

$$H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) + o_t(\hat{u}_t) \geq H_t(\hat{x}_t, u_t, \hat{\psi}_t) + o_t(u_t)$$

Teniendo en cuenta que la positividad de la función o_t y que $o_t(\hat{u}_t) = 0$. Deducimos que para toda $u_t \in U_t$

$$H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) + 0 \geq H_t(\hat{x}_t, u_t, \hat{\psi}_t) + o_t(u_t) \geq H_t(\hat{x}_t, u_t, \hat{\psi}_t).$$

Por lo tanto para $t = 0, \cdots, T-1$, $\hat{u}_t$ es solución del problema

$$\max_{u_t \in U_t} [H_t(\hat{x}_t, u_t, \hat{\psi}_t)],$$

donde $\hat{\psi}_t := \nabla_x g_{t+1}(\hat{x}_{t+1})$, por lo tanto $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ satisface el PFP. ■

Corolario 4.1.1 *El Teorema 4.1.1 es una consecuencia de la Proposición 4.1.6.*

Prueba. En efecto no es difícil ver de las hipótesis de afinidad respecto a la variable "x", que las funciones o_t definidas como en el Lema 4.1.3 satisfacen para toda $u_t \in U_t$

$$o_t(u_t) = 0.$$

■

Teniendo en cuenta el corolario anterior es posible generalizar el Teorema 4.1.1, obteniéndose el siguiente resultado.

Teorema 4.1.2 (PFP - Afin Convexo) *Consideremos el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo las hipótesis del Teorema 4.1.1 excepto la hipótesis $(H)_A$, en lugar de la cual satisface las siguientes:*

- *Para $t = 0, \cdots, T$, las funciones f_t^0 son convexas respecto a "x".*
- *Para $t = 0, \cdots, T-1$, las funciones f_t^0 son son afines respecto a "x".*

Sea $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ una solución óptima del problema, entonces satisface el PFP (es una sucesión PFP).

Prueba. Para la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$, y su proceso asociado $(\hat{x}_0, \cdots, \hat{x}_{T-1})$, consideremos las funciones $o_t : U_t \to \mathbb{R}$,

$$o_t(u_t) := g_{t+1}(f_t(\hat{x}_t, u_t)) - g_{t+1}(f_t(\hat{x}_t, \hat{u}_t)) - \langle \hat{\psi}(t); (f_t(\hat{x}_t, u_t) - f_t(\hat{x}_t, \hat{u}_t)) \rangle.$$

Por la proposición anterior para demostrar este teorema bastaría probar que $o_t(u_t) \geq 0$, para toda $u_t \in U_t$.

Teniendo en cuenta que $\hat{\psi}(t) = \nabla_x g_{t+1}(\hat{x}_{t+1}) = \nabla_x g_{t+1}(f_t(\hat{x}_t, \hat{u}_t))$, y definiendo $y := f_t(\hat{x}_t, u_t)$, la positividad de $o_t(u_t)$, es equivalente a la siguiente desigualdad

$$g_{t+1}(y) - g_{t+1}(\hat{x}_{t+1}) \geq \langle \nabla g_{t+1}(\hat{x}_{t+1}); y - \hat{x}_{t+1} \rangle. \qquad (\phi)$$

Lo cual sería cierto si la función g_{t+1} sería convexa. Vamos a demostrar que bajo las hipótesis de este teorema se deduce que las funciones g_t son convexas. Probemoslo por inducción. Primero $g_T := f_T^0$ es convexa por hipótesis.

Por inducción supongamos que hemos probado que la función g_{r+1} es convexa. Entonces de la definición de la función g_r tenemos

$$g_r(x_r) = f_r^0(x_r, \hat{u}_r) + g_{r+1}(f_r(x_r, \hat{u}_r)).$$

Como ambos sumandos son convexos por hipótesis se sigue que g_r es también convexo.

Por lo tanto para $t = 0, \cdots, T-1$, la función g_t es convexa, de lo que deducimos que la desigualdad (ϕ) es cierta. ■

Observación El teorema anterior generaliza al Teorema 4.1.1 pues las hipótesis que reemplazan a $(H)_A$, evidentemente la generalizan.

Es difícil obtener una generalización mayor en la que no se pida que las funciones g_t sean convexas. Es por ello que si quisiesemos generalizar más el teorema debemos buscar que las funciones f_t satisfagan ciertas hipótesis que permitan que con la composición de funciones se mantenga la convexidad, en este sentido podemos obtener tal generalización para el problema $P_T(\hat{x}_0)$ en el que los estados "x_t" son números reales ($x_t \in \mathbb{R}$).

Proposición 4.1.7 (PFP - Convexidad Creciente) *Consideremos el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo las hipótesis del Teorema 4.1.1 excepto la hipótesis $(H)_A$, en lugar de la cual satisface las siguientes:*

- *Para $t = 0, \cdots, T$, las funciones f_t^0 son convexas crecientes en "x".*
- *Para $t = 0, \cdots, T-1$, las funciones f_t^0 son convexas crecientes en "x".*
- *Para $t = 0, \cdots, T$, se tiene $X_t \in \mathbb{R}$.*

Sea $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ una solución óptima del problema, entonces satisface el PFP (es una sucesión PFP).

Prueba. La demostración es idéntica a la anterior, se reduce a demostrar que las funciones g_t asociadas a la sucesión $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ sean convexas.

Nuevamente probemos esto inductivamente. Primero $g_T := f_T^0$ es convexa por hipótesis.

Por inducción supongamos que hemos probado que la función g_{r+1} es convexa y creciente. Entonces de la definición de la función g_r tenemos

$$g_r(x_r) = f_r^0(x_r, \hat{u}_r) + g_{r+1}(f_r(x_r, \hat{u}_r)).$$

Como por hipótesis ambos sumandos son convexos y crecientes (la propiedad de ser convexa y creciente se mantiene bajo la composición) se sigue que g_r es también convexo y creciente. ■

Observación. La proposición anterior exige que los estados x_t sean números reales pues se necesita que las funciones f_t sean "crecientes", es decir se requiere un orden numérico. Si conseguiríamos dar un "orden" en espacios $\mathbb{R}^n$, existe la posiblidad de generalizar el resultado anterior para ciertos problemas de control, con estados en $\mathbb{R}^n$.

El teorema anterior generaliza en ciertos aspectos al Teorema PFP-afín (Teorema 4.1.1), pues la hipótesis de afinidad se debilitan con las de convexidad creciente.

Con este resultado finalizamos esta sección y dirigiremos la investigación en otra dirección. Mantenemos la pregunta inicial ¿qué condiciones debe cumplir el problema $\mathcal{P}_T(\hat{x}_0)$ para que el PFP sea una condición necesaria de optimalidad? Vamos a estudiar otra aproximación a la respuesta de esta pregunta, desarrollada por el matemático Philipe Michel, el cual uso la Programación Mixta para dar estas condiciones. Veremos además que esta aproximación es independiente a la realizada en esta sección. Es decir que existen problemas afines en "x" que no son mixtos, y viceversa. Por lo tanto no se puede deducir del Teorema 4.1.1 las conclusiones de la siguiente sección y viceversa.

4.2 Programación Mixta

El matemático francés Philipe Michel estudió un tipo de problemas de optimización estática, denominados **programas mixtos** [19]. Usando éstos es posible estudiar un cierto tipo de problemas de control que satisface el Principio Fuerte de Pontryagin.

En esta sección estudiaremos con cuidado la definición de estos problemas así como su relación con los problemas de control óptimo de procesos discretos.

4.2.1 Programas Mixtos

Para definir los problemas de la Programación Mixta, es decir los problemas mixtos, nos restringuiremos a trabajar con pr oblemas del tipo siguiente,

$$\begin{array}{rll} & \max & f^0(x,u) \\ i = 1, \cdots, m_0, & f^i(x,u) \geq 0, & \\ i = m_0+1, \cdots, m, & f^i(x,u) = 0, & \end{array} \qquad (\mathcal{P})$$

donde $f^i : X \times U \to \mathbb{R}$, $X \subseteq \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$. Se exigirá que se cumpla dos condiciones:

- X es un conjunto convexo,
- Para cada "u" fijo la función, $f^i(\cdot, u) : X \to \mathbb{R}$ es diferenciable. Con $i = 1, \cdots m_0, \cdots, m$.

A lo largo de esta sección siempre se trabajará con problemas $(\mathcal{P})$ **que satisfagan estas dos condiciones.**

Definición 4.2.1 (Programa Mixto) *Se dice que el problema* $\mathcal{P}$ *es* **mixto en** $(\bar{x}, \bar{u})$ *si la siguiente condición es verdadera:*

Si $(\bar{x},\bar{u})$ es una solución óptima de $\mathcal{P}$, entonces para todo conjunto finito $\{u_1,\cdots,u_N\}$ de U, $(\bar{x},0)$ es solución óptima del problema $\mathcal{G}(u_1,\cdots,u_N)$ definido como:

$$\begin{array}{lll} & \max \quad g^0(x,u) & \\ i=1,\cdots,m_0, & g^i(x,u)\geq 0, & \mathcal{G}(u_1,\cdots,u_N) \\ i=m_0+1,\cdots,m & g^i(x,u)=0. & \end{array}$$

Donde para $i=0,\cdots,m$,

$$\begin{array}{rccl} g^i: & X\times Y & \to & \mathbb{R} \\ & (x,y) & \mapsto & \sum_{j=1}^N y^j f^i(x,u_j)+(1-\sum_{j=1}^N y^j)f^i(x,\bar{u}), \end{array}$$

con $Y:=\{(y^1,\cdots,y^N)\in\mathbb{R}^N/y^j\geq 0,\sum_{j=1}^N y^j\leq 1$ para $j=1,\cdots,N\}$.

Esta condición se resume en: $(\bar{x},0)$ es solución óptima de cada problema de la familia de problemas,

$$\mathcal{G}:=\{\ \mathcal{G}(u_1,\cdots,u_N):\{u_1,\cdots,u_N\}\subseteq U,N\in\mathbb{N}\ \}.$$

Por último se dirá que el problema $(\mathcal{P})$ es **mixto en todo punto** *si $(\mathcal{P})$ es mixto para todo $(x,u)\in X\times U$.*

Observación. Usando la lógica de proposiciones, la condición de la definición será verdadera para todo (x,u) que no sea óptimo de $(\mathcal{P})$ (pues $F\to V$ y $F\to F$ son verdaderos). Por lo tanto $(\mathcal{P})$ es mixto para todo (x,u) no óptimo. En particular si $(\mathcal{P})$ no tiene solución entonces es mixto en todo punto.

De esto se deduce además que para saber si un problema $(\mathcal{P})$ es mixto en todo punto, basta estudiar si es mixto en sus soluciones óptimas (si existen). Este hecho será muy usado más adelante, por ello es preciso tenerlo en cuenta.

Estos aspectos de la definición parecen ser innecesarios pero al no implicar contradicción alguna no cambiaremos la definición de los programas mixtos.

Finalmente podemos interpretar esta definición diciendo que los problemas mixtos vienen a ser aquellos en los que el óptimo se "conserva" para ciertas "variaciones convexas".

Dado un problema concreto, ¿Cómo podemos saber si es mixto? En esta dirección tenemos una condición suficiente y otra que es necesaria. La primera fue desarrollada por Philip Michel.

Condición Suficiente Recordemos que estamos trabajando con el problema $(\mathcal{P})$. Para $x\in X$, definimos los siguientes conjuntos

$$A(x)=\left\{ z=(z^0,z^1,\cdots,z^m)\in\mathbb{R}^{m+1} \;\middle|\; \begin{array}{l} \text{Existe} \quad u\in U \quad \text{tal que} \\ z^i\leq f^i(x,u), \quad i=0,\cdots,m_0, \\ z^i=f^i(x,u), \quad i=m_0+1,\cdots,m \end{array} \right\}.$$

$$B(x)=\left\{ z=(z^0,z^1,\cdots,z^m)\in\mathbb{R}^{m+1} \;\middle|\; \begin{array}{l} \text{Existen } \; u\in U \text{ y } v\in\mathbb{R}^m \text{ tales que} \\ \quad z^0\le f^0(x,u), \\ v^i.z^i\le f^i(x,u), \quad i=0,\cdots,m_0, \\ v^i.z^i=f^i(x,u), \quad i=m_0+1,\cdots,m \end{array} \right\}.$$

No es difícil demostrar desde las definiciones que siempre se cumple

$$A(x)\subseteq B(x).$$

De la relación de estos conjuntos obtenemos el siguiente criterio (condición suficiente) para que un problema $(\mathcal{P})$ sea mixto.

Lema 4.2.1 (Condición Suficiente) *Consideremos el problema* $(\mathcal{P})$*. Si para todo* $x\in X$ *se tiene* $\mathrm{conv}A(x)\subseteq B(x)$*. Entonces* $(\mathcal{P})$ *es un problema* **mixto en todo punto.**

Prueba. Si el problema no tiene solución entonces vimos en la observación que sigue a la definición, que $\mathcal{P}$ es mixto en todo punto. Luego, no hay nada que probar en este caso. De otra parte si el problema tiene al menos una solución, probemos por contradicción que el problema es mixto en cada una de ellas.

Sea $(\bar{x},\bar{u})$ una solución de $\mathcal{P}$ fija y arbitraria. Supongamos que exista un conjunto $\{u_1,\cdots,u_N\}\subset U$, tal que $(\bar{x},0)$ no es solución óptima del problema $\mathcal{G}(u_1,\cdots,u_N)$.

Entonces existe $(x,y)\in X\times Y$ tal que,

$$g^0(x,y)>g^0(\bar{x},0),$$

satisfaciendo

$$\begin{array}{ll} g^i(x,y)\ge 0, & i=1,\cdots,m_0, \\ g^i(x,y)=0, & i=m_0+1,\cdots,m. \end{array}$$

Para cada u_j vamos a formar un elemento $z_j\in A(x)$, definiendo sus coordenadas,

$$\begin{array}{lll} z_j^0:=f^0(x,u_j) & \le f^0(x,u_j), & \\ z_j^i:=f^i(x,u_j)-g^i(x,y) & \le f^i(x,u_j), & \text{para } i=1,\cdots,m_0, \\ z_j^i:=f^i(x,u_j)-g^i(x,y) & = f^i(x,u_j), & \text{para } i=m_0+1,\cdots,m. \end{array}$$

Es claro que para $j=0,\cdots,N$: $z_j:=(z_j^0,z_j^1,\cdots,z_j^{m_0},z_j^{m_0+1},\cdots,z_j^m)\in A(x)$. Entonces

$$z:=\sum_{j=1}^N y^j z_j+(1-\sum_{j=1}^N y^j)z_0\in conv(A(x)).$$

Por hipótesis conv$A(x) \subseteq B(x)$, por ello $z \in B(x)$. Calculemos los componentes de $z = (z^0, z^1, \cdots, z^m)$,

$$z^0 = \sum_{j=1}^{N} y^j z_j^0 + (1 - \sum_{j=1}^{N} y^j) z_0^0 = \sum_{j=1}^{N} y^j f^0(x, u_j) + (1 - \sum_{j=1}^{N} y^j) f^0(x, \bar{u}) = g^0(x, y),$$

para $1 \leq i \leq m$: $z^i = 0$. En efecto calculemos

$$\begin{array}{rcl} z^i & = & \sum_{j=1}^{N} y^j z_j^i + (1 - \sum_{j=1}^{N} y^j) z_0^i \\ & = & \sum_{j=1}^{N} y^j (f^i(x, u_j) - g^i(x, y)) + (1 - \sum_{j=1}^{N} y^j)(f^i(x, \bar{u}) - g^i(x, y)) \\ & = & \sum_{j=1}^{N} y^j f^i(x, u_j) - g^i(x, y) + (1 - \sum_{j=1}^{N} y^j) f^i(x, \bar{u}) \\ & = & g^i(x, y) - g^i(x, y) = 0. \end{array}$$

por lo tanto $z = (g^0(x, y), 0, \cdots, 0) \in B(x)$. De la definición de $B(x)$ tenemos que existe $u \in U$ y $v \in \mathbb{R}^m$ tales que:

$$\begin{array}{rcll} z_0 & \leq & f^0(x, u), & \\ v^i z^i = 0 & \leq & f^i(x, u), & \text{para } i = 1, \cdots, m_0, \\ v^i z^i = 0 & = & f^i(x, u), & \text{para } i = m_0 + 1, \cdots, m. \end{array}$$

De esto se deduce que (x, u) es admisible para $\mathcal{P}$ y $z_0 \leq f^0(x, u)$.

Desde que $z_0 = g^0(x, y) > g^0(\bar{x}, 0) = f^0(\bar{x}, \bar{u})$ obtenemos (x, u) admisible tal que $f^0(x, u) > f^0(\bar{x}, \bar{u})$. Por lo tanto $(\bar{x}, \bar{u})$ no es solución óptima del problema $(\mathcal{P})$. Esto es una contradicción que nace de suponer que $(\bar{x}, 0)$ no es solución del problema $\mathcal{G}(u_1, \cdots, u_N)$. Por lo tanto hemos demostrado que $(\bar{x}, 0)$ es solución de cualquier problema de la familia $\mathcal{G}$. Así pues en vista de que $(\bar{x}, \bar{u})$ es una solución arbitraria entonces $\mathcal{P}$ es mixto para todo punto. ∎

Recordemos que a partir de las definiciones de los conjuntos $A(x)$ y $B(x)$ siempre es cierto que $A(x) \subseteq B(x)$. Teniendo en cuenta este hecho junto con el Lema 4.2.1, obtenemos el siguiente corolario.

Corolario 4.2.1 *Para el problema* $(\mathcal{P})$. *Si para todo* $x \in X$ *se tiene* $A(x)$ *es convexo o* $B(x)$ *es convexo. Entonces el problema* $(\mathcal{P})$ *es* **mixto en todo punto**.

Prueba. Si $A(x)$ es convexo entonces para $x \in X$, conv$A(x) = A(x) \subseteq B(x)$. Por lo tanto del Lema 4.2.1, $\mathcal{P}$ es mixto en todo punto.

De otra parte si $B(x)$ es convexo, entonces desde que para todo $x \in X$, $A(x) \subseteq B(x)$, se deduce que conv$A(x) \subseteq B(x)$. Entonces del Lema 4.2.1 se deduce lo que queremos. ∎

Usando la definición del conjunto $A(x)$ vemos que unas condiciones para que sea convexo son las siguientes.

Corolario 4.2.2 *Consideremos el problema* $(\mathcal{P})$ *satisfaciendo.*

- U *es convexo.*
- f^0 *es convexo con respecto a la variable* u.
- $f^1, \cdots, f^{m_0}$, *son cóncavos respecto a la variable* u.
- $f^{m_0+1}, \cdots, f^n$ *son lineal afines respecto a la variable* u.

Entonces el problema $(\mathcal{P})$ *es un problema mixto* **para todo punto**.

Prueba. Del Lema 4.2.1 vemos que sólo es necesario demostrar que para todo $x \in X$. El conjunto $A(x)$ es convexo. Tomemos $\bar{z}, \hat{z} \in A(x)$ fijos y arbitrarios Para $\alpha \in [0,1]$ vamos a demostrar que $\alpha\bar{z} + (1-\alpha)\hat{z} \in A(x)$.

De la definición del conjunto $A(x)$, existen $\bar{u}$ y $\hat{u}$ tales que:

$$\begin{array}{lll} \bar{z}^i \leq f^i(x,\bar{u}), & \hat{z}^i \leq f^i(x,\hat{u}), & i = 0, \cdots, m_0, \\ \bar{z}^i = f^i(x,\bar{u}), & \hat{z}^i = f^i(x,\hat{u}), & i = m_0+1, \cdots, m. \end{array}$$

Como U es convexo tenemo que $\tilde{u} := \alpha\bar{u} + (1-\alpha)\hat{u} \in U$. No es difícil demostrar desde las hipótesis que es cierto que $\alpha\bar{z} + (1-\alpha)\hat{z}$

$$\begin{array}{ll} \alpha\bar{z} + (1-\alpha)\hat{z} \leq f^i(x,\tilde{u}), & i = 0, \cdots, m_0, \\ \alpha\bar{z} + (1-\alpha)\hat{z} = f^i(x,\tilde{u}), & i = m_0+1, \cdots, m. \end{array}$$

Por lo tanto probamos el corolario. ■

Habiendo demostrado esta condición suficiente, nos preguntamos si es necesaria también. La respuesta es negativa. Existen problemas mixtos que no satisfacen tal condición.

Contraejemplo: Condición Necesaria. Consideremos el problema

$$\max_{f^1(x,u)=0} f^0(x,u), \tag{$\mathcal{P}$}$$

donde

$$\begin{array}{rlcl} f^0: & X \times U & \to & \mathbb{R} \\ & (x,u) & \mapsto & u^2, \end{array} \qquad \begin{array}{rlcl} f^1: & X \times U & \to & \mathbb{R} \\ & (x,u) & \mapsto & x-u, \end{array}$$

con $X := \mathbb{R}$, $U := [-c,c]$ y $0 < c \in \mathbb{R}$.

Es claro que las soluciones óptimas de este problema son solamente dos: $(\bar{x},\bar{u}) := (c,c)$ y $(\hat{x},\hat{u}) := (-c,-c)$. Probemos que $\mathcal{P}$ es mixto en (c,c). En efecto consideremos un conjunto finito $\{u_1, \cdots, u_N\} \subseteq U$ fijo y arbitrario. Formemos el problema $\mathcal{G}(u_1, \cdots, u_N)$,

$$\max_{g^1(x,y)=0} g^0(x,y).$$

Tomemos (x,y) fijo y arbitrario y admisible. Entonces por su definición

$$g^0(x,y)=\sum_{j=1}^{N} y^j f^i(x,u_j)+(1-\sum_{j=1}^{N} y^j)f^i(x,\bar{u})\leq \sum_{j=1}^{N} y^j c^2+(1-\sum_{j=1}^{N} y^j)c^2\leq c^2.$$

Por lo cual para todo (x,y) admisible $g^0(x,y)\leq c^2=g^0(\bar{x},0)$. Así pues $(\bar{x},0)=(c,0)$ es óptimo del problema $\mathcal{G}(u_1,\cdots,u_N)$.

Con similar demostración obtenemos que $\mathcal{P}$ es mixto en $(-c,-c)$. Por lo tanto $\mathcal{P}$ es mixto en todo punto.

Demostremos ahora que no satisface la condición suficiente. Concretamente para $x=c/2$,

$$conv(A(c/2))\not\subseteq B(c/2).$$

En efecto tomese

$$\begin{array}{ll} \hat{z}:=(\hat{a}=c^2,\ \hat{b}:=-\frac{c}{2})\in A(c/2) & \text{(tomese } \hat{u}=\ c),\\ \bar{z}:=(\bar{a}=c^2,\bar{b}:=-\frac{3c}{2})\in A(c/2) & \text{(tomese } \bar{u}=-c). \end{array}$$

Para $\alpha=1/4$, tenemos

$$(c^2,0)=(\hat{a},\hat{b})+\frac{1}{4}[(\bar{a},\bar{b})-(\hat{a},\hat{b})]\in convA(c/2).$$

Supongamos que $(c^2,0)\in B(c/2)$, entonces existe $u\in[-c,c]$ y $v\in\mathbb{R}$ tales que:

$$\begin{array}{c} c^2\leq f^0(c/2,u)=u^2,\\ v.0=0=f^1(c/2,u)=c/2-u. \end{array}$$

Lo cual es imposible. Por ello $(c^2,0)\notin B(c/2)$. ■

Usando el método de demostración de que el problema anterior es mixto en todo punto obtenemos la siguiente condición suficiente para ser problema mixto en todo punto.

Lema 4.2.2 (2da. Condición Suficiente) *Para el problema $\mathcal{P}$ supongamos que $f^0(x,u)=f^0(u)$ y además*

- *Existe $\bar{u}\in U$ tal que para toda $u\in U$: $f^0(\bar{u})\geq f^0(u)$.*
- *Existe $\bar{x}\in X$ tal que $(\bar{x},\bar{u})$ es admisible para el problema.*

Entonces $\mathcal{P}$ es mixto en todo punto.

Prueba. Sabemos que sólo es necesario demostrar que $\mathcal{P}$ es mixto para las soluciones óptimas (si es que existen). Si no existiesen entonces automáticamente es mixto en todo punto. Por ello sea $(\hat{x},\hat{u})$ una solución fija y arbitraria. Entonces

$f^0(\hat{u}) = f^0(\hat{x}, \hat{u}) \geq f^0(\bar{x}, \bar{u}) = f^0(\bar{u})$. Por hipótesis $f^0(\bar{u}) \geq f^0(\hat{u})$, se deduce que $f^0(\hat{u}) = f^0(\hat{u})$.

Consideremos ahora un conjunto fijo y arbitrario $\{u_1, \cdots, u_N\} \subseteq U$ y el problema asociado: $\mathcal{G}(u_1, \cdots, u_N)$. Tomemos $(x, y) \in X \times Y$ fijo y arbitrario. Se cumple

$$\begin{array}{rcl} g^0(x, y) & = & \sum_{j=1}^N y^j f^0(x, u_j) + (1 - \sum_{j=1}^N y^j) f^0(x, \hat{u}) \\ & = & \sum_{j=1}^N y^j f^0(u_j) + (1 - \sum_{j=1}^N y^j) f^0(\hat{u}) \\ & \leq & f^0(\bar{u}) = f^0(\hat{u}) \\ & = & g^0(\hat{x}, 0). \end{array}$$

Como el problema $\mathcal{G}(u_1, \cdots, u_N)$ era arbitrario, deducímos que $(\hat{x}, 0)$ es solución para cada uno de los problemas de la familia $\mathcal{G}$. Por ello $\mathcal{P}$ es mixto en $(\hat{x}, \hat{u})$. Como este es una solución arbitraria, deducimos que $\mathcal{P}$ es mixto en todo punto. ■

Ahora pasemos a estudiar una condición necesaria para que $\mathcal{P}$ sea mixto en un punto $(\bar{x}, \bar{u})$. Para ello definamos el siguiente conjunto asociado a $(\bar{x}, \bar{u})$

$$\mathcal{U}(\bar{x}) = \left\{ u \in U \;\middle|\; \begin{array}{cc} |f^i(\bar{x}, u)| \leq |f^i(\bar{x}, \bar{u})|, & \text{para } i = 1, \cdots, m_0, \\ f^i(\bar{x}, u) = 0, & \text{para } i = m_0 + 1, \cdots, m. \end{array} \right\}.$$

Lema 4.2.3 (Condición Necesaria) *En el problema $\mathcal{P}$, sea $(\bar{x}, \bar{u})$ una solución óptima. Si $\mathcal{P}$ es mixto en $(\bar{x}, \bar{u})$, entonces $\bar{u}$ es solución del problema*

$$\max_{u \in \mathcal{U}(\bar{x})} f^0(\bar{x}, u).$$

Prueba. Es claro que $\bar{u} \in \mathcal{U}(\bar{x})$, por lo cual $\mathcal{U}(\bar{x}) \neq \emptyset$. Consideremos $u \in \mathcal{U}(\bar{x})$ fijo y arbitrario. Desde que $\mathcal{P}$ es mixto en $(\bar{x}, \bar{u})$ entonces $(\bar{x}, 0)$ es solución del problema $\mathcal{G}(u)$.

Usando las hipótesis vamos a demostrar que existe $y > 0$ suficientemente pequeño tal que $(\bar{x}, y) \in \operatorname{Adm} \mathcal{G}(\mathrm{u})$.

En efecto definamos el conjunto

$$J^c(u) = \{i = 1, \cdots, m_0 : |f^i(\bar{x}, u)| < f^i(\bar{x}, \bar{u})\}.$$

Supongamos que $J^c(u) = \emptyset$, entonces $f^i(\bar{x}, u) = f^i(\bar{x}, \bar{u})$. (para $i = 1, \cdots, m_0$).

Tomemos $y = 1/2$, entonces usando la definición de las funciones g^i y $\mathcal{U}(\bar{x})$ tenemos

$$\begin{array}{rcll} 1/2 f^i(\bar{x}, u) + (1/2) f^i(\bar{x}, \bar{u}) & = & f^i(\bar{x}, \bar{u}) \geq 0 & \text{para } i = 1, \cdots, m_0, \\ 1/2 f^i(\bar{x}, u) + (1/2) f^i(\bar{x}, \bar{u}) & = & f^i(\bar{x}, \bar{u}) = 0 & \text{para } i = m_0 + 1, \cdots, m. \end{array}$$

Por lo tanto $(\bar{x}, 1/2) \in \operatorname{Adm} \mathcal{G}(\mathrm{u})$.

Si por el contrario $J^c(u) \neq \emptyset$, entonces tomemos $y \in \mathbb{R}$ tal que

$$0 < y \leq \frac{1}{2} \min \left\{ \frac{f^i(\bar{x},\bar{u})}{f^i(\bar{x},\bar{u}) - f^i(\bar{x},u)} \; : \; i \in J^c(u) \right\}.$$

Probemos que $(\bar{x}, y) \in \operatorname{Adm} \mathcal{G}(\mathrm{u})$. En efecto para $j \in J^c(u)$,

$$\begin{array}{rcl} g^j(\bar{x},y) & = & f^j(\bar{x},\bar{u}) - y[f^j(\bar{x},\bar{u}) - f^j(\bar{x},u)] \\ & \geq & f^j(\bar{x},\bar{u}) - \frac{1}{2}\min\left\{\frac{f^i(\bar{x},\bar{u})}{f^i(\bar{x},\bar{u})-f^i(\bar{x},u)}\right\}[f^j(\bar{x},\bar{u}) - f^j(\bar{x},u)] \\ & \geq & f^j(\bar{x},\bar{u}) - \frac{1}{2}\{\frac{f^j(\bar{x},\bar{u})}{f^j(\bar{x},\bar{u})-f^j(\bar{x},u)}[f^j(\bar{x},\bar{u}) - f^j(\bar{x},u)]\} \\ & = & \frac{1}{2}f^j(\bar{x},\bar{u}) \geq 0. \end{array}$$

Para $j \in \{1, \cdots, m_0\} \setminus J^c(u)$, tenemos que $f^j(\bar{x},\bar{u}) = f^j(\bar{x},u)$. Por lo cual $g^j(\bar{x},y) = yf^j(\bar{x},u) + (1-y)f^j(\bar{x},\bar{u}) = f^j(\bar{x},\bar{u}) \geq 0$.

Por otra parte para $j = m_0+1, \cdots, m$: $0 = f^j(\bar{x},\bar{u}) = f^j(\bar{x},u)$ entonces es claro que $g^j(\bar{x},y) = 0$. Por lo cual $(\bar{x},y) \in \operatorname{Adm} \mathcal{G}(\mathrm{u})$.

Finalmente de la optimalidad de $(\bar{x},0)$ deducimos secuencialmente

$$\begin{array}{l} g^0(\bar{x},0) \geq g^0(\bar{x},y), \\ f^0(\bar{x},\bar{u}) \geq yf^0(\bar{x},u) + (1-y)f^0(\bar{x},\bar{u}), \\ yf^0(\bar{x},\bar{u}) \geq yf^0(\bar{x},u), \\ f^0(\bar{x},\bar{u}) \geq f^0(\bar{x},u). \end{array}$$

Como $u \in \mathcal{G}(\bar{x})$ es arbitrario de la desigualdad anterior obtenemos que $\bar{u}$ es solución del problema

$$\max_{u \in \mathcal{U}(\bar{x})} f^0(\bar{x},\bar{u}).$$

■

Observación. No creemos que la condición necesaria anterior sea también suficiente, debido a que el conjunto $\mathcal{U}(\bar{x})$ es demasiado limitado y no parece lo suficiente grande como para asegurar que $(\bar{x},0)$ sea óptimo de cualquier problema $\mathcal{G}(u_1, \cdots, u_N)$. Buscar un contraejemplo explícito es un problema abierto. En vista que nuestro interés principal no va en esta dirección no continuaremos por este camino.

Por otra parte para los problemas mixtos la condición necesaria de óptimo toma una forma especial.

Teorema 4.2.1 (Condición Necesaria de Optimalidad) *Consideremos en el problema $(\mathcal{P})$ una solución óptima $(\bar{x},\bar{u})$. Si $(\mathcal{P})$ es mixto en $(\bar{x},\bar{u})$. Entonces existe $a := (a_0, \cdots, a_m) \in \mathbb{R}^{m+1}$, con $i = 0, \cdots, m$, tales que:*

1. $\sum_{i=0}^m |a_i| = 1$.

2. *Para* $i = 1, \cdots, m_0$, $\qquad a_0 \geq 0$, $a_i \leq 0$.

3. *Para* $i = 1, \cdots, m_0$, $\quad a_i f^i(\bar{x}, \bar{u}) = 0.$

4. *Definiendo* $L(x, u, a) = \sum_{i=0}^{m} a_i f^i(x, u)$*, se cumple: Para todo* $x \in X$

$$\sum_{k=1}^{n} (x^k - \bar{x}^k) \frac{\partial L}{\partial x^k}(\bar{x}, \bar{u}, a) \leq 0.$$

5. $\bar{u}$ *es solución del problema*

$$\max_{u \in U} L(\bar{x}, u, a).$$

Prueba. La prueba tiene dos partes. En la primera se usa la condición necesaria de Fritz John en un problema de la familia $(\mathcal{G})$ fijo y arbitrario. Seguidamente se usa el Teorema de Intersección Finita (PIF) para concluir el resultado final.

Por hipótesis $(\mathcal{P})$ es mixto en $(\bar{x}, \bar{u})$ entonces $(\bar{x}, 0)$ es solución de cualquier problema de la familia $\mathcal{G}$. Tomemos un problema fijo y arbitrario $\mathcal{G}(u_1, \cdots, u_N)$ el cual es equivalente a

$$\mathcal{G}(u_1, \cdots, u_N) \equiv \left| \begin{array}{lll} \quad \min & -g^0(x, u) & \\ -g^i(x, u) \leq 0, & & i = 1, \cdots, m_0 \\ \ g^i(x, u) = 0, & & i = m_0 + 1, \cdots, m. \end{array} \right.$$

Por lo tanto $(\bar{x}, 0)$ es óptimo de este último problema. Aplicamos la condición de Fritz John tenemos que existen números $\tilde{a}_0, \tilde{a}_1, \cdots, \tilde{a}_m \in \mathbb{R}$ **no todos nulos** tales que:

(1)

$$\sum_{j=0}^{m_0} \tilde{a}_j \nabla_{(x,y)}(-g^j(\bar{x}, 0)) + \sum_{j=m_0+1}^{m} \tilde{a}_j \nabla_{(x,y)} g^j(\bar{x}, 0) = 0$$

(2) Para $i = 0, \cdots, m_0$, $\quad \tilde{a}_i \geq 0.$

(3) Para $i = 1, \cdots, m_0$, $\quad \tilde{a}_i(g^i(\bar{x}, 0)) = 0.$

Definiendo

$$\left| \begin{array}{ll} a_0 := \tilde{a}_0, & \\ a_i := -\tilde{a}_i, & i = 1, \cdots, m_0, \\ a_i := \tilde{a}_i, & i = m_0 + 1, \cdots, m. \end{array} \right.$$

Con esto podemos decir que existen $a_0, \cdots, a_m \in \mathbb{R}$ tales que:

- $a_0(-\nabla_{(x,y)} g^0(\bar{x}, 0)) + \sum_{i=1}^{m} a_i \nabla_{(x,y)} g^i(\bar{x}, 0) = 0 \in \mathbb{R}^n \times \mathbb{R}^N$,
- $a_i g^i(\bar{x}, 0) = 0$, $\quad$ para $\quad i = 1, \cdots, m_0$,
- $a_0 \geq 0$, $a_i \leq 0$, $\quad$ para $\quad i = 1, \cdots, m_0$.

De otro lado como $X \times Y$ es convexo entonces para todo $(x, y) \in X \times Y$ la dirección $d := (x, y) - (\bar{x}, 0)$ es factible. Desde la optimalidad de $(\bar{x}, 0)$ y de la Condición Necesaria de Optimalidad Primal (ver [15, p.44]) tenemos

$$\langle \nabla_{(x,y)} g^0(\bar{x}, 0); (x - \bar{x}, y) \rangle \leq 0.$$

Observando que $a_0 \geq 0$ y desde **(1)**, $a_0 g^0(\bar{x}, 0) = \sum_{i=1}^m a_i \nabla_{(x,y)} g^i(\bar{x}, 0)$ obtenemos

$$\left\langle \sum_{i=1}^m a_i \nabla_{(x,y)} g^i(\bar{x}, 0) \; ; \; (x - \bar{x}, y) \right\rangle \leq 0.$$

Sumando ambas desigualdades

$$\left\langle a_0 \nabla_{(x,y)} g^0(\bar{x}, 0) + \sum_{i=1}^m a_i \nabla_{(x,y)} g^i(\bar{x}, 0) \; ; \; (x - \bar{x}, y) \right\rangle \leq 0.$$

Usando las coordenadas de las variables $(x, y) = (x_1, \cdots, x_n, y_1, \cdots y_N)$ en el cálculo de gradientes

$$\sum_{k=1}^n (x^k - \bar{x}^k) \left[\sum_{i=0}^m a_i \frac{\partial}{\partial x^k} g^i(\bar{x}, 0) \right] + \sum_{j=1}^N y^j \left[\sum_{i=0}^m a_i \frac{\partial}{\partial y^j} g^i(\bar{x}, 0) \right] \leq 0. \qquad (\alpha)$$

Lo anterior vale para todo elemento $(x, y) \in X \times Y$, en particular para elementos del tipo $(x, 0)$, por ello obtenemos que para todo $x \in X$

$$\sum_{k=1}^n (x^k - \bar{x}^k) \left[\sum_{i=1}^m a_i \frac{\partial}{\partial x^k} g^i(\bar{x}, 0) + a_0 \frac{\partial}{\partial x^k} g^0(\bar{x}, 0) \right] \leq 0.$$

Expresándolo usando el Lagrangiano obtenemos **[4.]**

$$\sum_{k=1}^n (x^k - \bar{x}^k) \left[\frac{\partial}{\partial x^k} L(\bar{x}, \bar{u}, a_0, a_1, \cdots, a_m) \right] = \sum_{k=1}^n (x^k - \bar{x}^k) \frac{\partial}{\partial x^k} L(\bar{x}, \bar{u}, a) \leq 0.$$

Desde que a_i no son todos nulos podemos normalizarlos $\sum_{j=0}^n |a_j| = 1$.

De otro lado nuevamente en la desigualdad (α) tomemos (x, y) de modo que $x = \bar{x}$ y $y^i = 0$ si $i \neq j$ y $y^j = 1$, obtenemos

$$a_0 \frac{\partial}{\partial y^j} g^0(\bar{x}, 0) + \sum_{i=1}^m a_i \frac{\partial}{\partial y^j} g^i(\bar{x}, 0) \leq 0. \qquad (\beta)$$

Desde la definición de g^i calculamos $\frac{\partial}{\partial y^j} g^i(\bar{x}, 0) = f^0(\bar{x}, u_j) - f^0(\bar{x}, \bar{u})$. Por lo tanto (β) es equivalente a

$$a_0(f^0(\bar{x}, u_j) - f^0(\bar{x}, \bar{u})) + \sum_{i=1}^m a_i(f^i(\bar{x}, u_j) - f^i(\bar{x}, \bar{u})) \leq 0.$$

Reordenando e interpretando con el lagrangiano L obtenemos para $j = 1, \cdots, N$, las siguientes desigualdades

$$L(\bar{x}, u_j, a_0, a_1, \cdots, a_m) \leq L(\bar{x}, \bar{u}_j, a_0, a_1, \cdots, a_m).$$

En vista de que el problema $\mathcal{G}(u_1, \cdots, u_N)$ fue elegido arbitrariamente obtenemos el siguiente resultado.
Para cada conjunto $Z := \{u_1, \cdots, u_N\} \subseteq U$ existe $a^Z = (a_0^Z, \cdots, a_m^Z) \in \mathbb{R}^{m+1}$ **no nulo** tal que:

(1) $\sum_{j=0}^{n} |a_j^Z| = 1$.

(2) $a_0^Z \geq 0$ y $a_j^Z \leq 0$ para $j = 1, \cdots, m_0$.

(3) $a_j^Z f^j(\bar{x}, \bar{u}) = 0$ para $j = 1, \cdots, m_0$.

(4) $\sum_{k=1}^{n} (x^k - \bar{x}^k) \frac{\partial L}{\partial x^k}(\bar{x}, \bar{u}, a^Z) \leq 0$.

(5) Para cada u_j^Z,
$$L(\bar{x}, u_j^Z, a^Z) \leq L(\bar{x}, \bar{u}, a^Z),$$
con $j = 1, \cdots, N$.

Definamos el siguiente conjunto:

$$K(Z) = \{a \in \mathbb{R}^{m+1} : \text{ satisface } (1), \cdots, (5)\}$$

Por el resultado anterior vemos que $a^Z \in K(Z)$ por lo tanto se tiene que $K(Z) \neq \emptyset$. Usando las 5 propiedades no es difícil demostrar que $K(Z)$ es un conjunto compacto.

Por ello cada conjunto finito Z define un conjunto compacto $K(Z)$. En consecuencia tenemos la familia de conjuntos $(K(Z))_{Z \subset U}$. Probemos que esta familia satisface la propiedad de la intersección finita (PIF).

Afirmación Dados "s" subconjuntos finitos Z_r de U, con $r = 1, \cdots, s$. Siempre se cumple

$$K(\cup_{r=1}^{r=s} Z_r) \subseteq \cap_{r=1}^{r=s} K(Z_r).$$

Prueba. Las propiedades $(1), \cdots, (4)$ no dependen de Z por ello un elemento $a \in K(\cup_{r=1}^{r=s} Z_r)$ las satisface. Sólo faltaría probar que cumple la propiedad (5) del conjunto $K(Z_r)$ para demostrar que pertenece a él (con $r = 1, \cdots, s$).

En efecto de la propiedad (5) para $K(\cup_{r=1}^{r=s} Z_r)$ se tiene que para $u_j \in \cup_{r=1}^{r=s} Z_r$

$$L(\bar{x}, u_j, a_0, a_1, \cdots, a_m) \leq L(\bar{x}, \bar{u}, a_0, a_1, \cdots, a_m).$$

En particular para $u_j \in Z_r$ se cumple la desigualdad, esto es precisamente la propiedad (5) de $K(Z_r)$, por lo tanto la satisface. ■

Como consecuencia de la afirmación anterior tenemos usando el teorema de la propiedad de intersección finita que existe $a = (a_0, \cdots, a_m) \in \cap_Z K(Z)$ **no nulo** tal que satisface:

(1) $\sum_{j=0}^{n} |a_j| = 0$,

(2) $a_0 \geq 0$ y $a_j \leq 0$ para todo $j = 1, \cdots, m_0$,

(3) $a_j f^j(\bar{x}, \bar{u}) = 0$ para $j = 1, \cdots, m_0$.

(4) $\sum_{k=1}^{n}(x^k - \bar{x}^k)[\frac{\partial}{\partial x^k} L(\bar{x}, \bar{u}, a)] \leq 0$

(5) $\bar{u}$ es solución del problema

$$\max_{u \in U} L(\bar{x}, u, a).$$

Con lo cual el teorema esta probado. ■

Ahora ya tenemos la condición necesaria de óptimo para problemas mixtos, los cuales son problemas de optimización estática. Con la ayuda de este teorema podemos demostrar el PFP para Programas Mixtos.

4.2.2 PFP: Programas Mixtos

De un modo similar a la metodología utilizada en el Capítulo 3, podemos estudiar el problema de control óptimo utilizando su equivalente estático. Al cual podemos aplicar el Teorema 4.2.1. Antes de hacer esto planteemos explícitamente algunos supuestos que haremos.

- Para $t = 0, \cdots, T-1$,

$$f_t^0 : A_t \times B_t \to \mathbb{R} \;, \qquad\qquad f_t : A_t \times B_t \to \mathbb{R}^n,$$

  donde A_t y B_t son conjuntos abiertos tales que $X_t \subset A_t$ y $X_t \subset B_t$.

- Los conjuntos X_t y U_t están definidos usando funciones según,

$$X_t \times U_t := \{(x, u) \in A_t \times B_t : g_t(x, u) \geq 0\},$$

  donde $g_t(x, u) : A_t \times B_t \to \mathbb{R}^{n_t}$.

Esta forma de definir los conjuntos X_t y U_t generaliza a la antigua forma usada en la sección 1 y 2 del Capítulo 3, pues si

$$X_t := \{x \in A_t : p_t(x) \leq 0\}, \qquad U_t := \{u \in B_t : q_t(x) \leq 0\},$$

con $p_t : A_t \to \mathbb{R}^{n_t}$ y $q_t : B_t \to \mathbb{R}^{m_t}$. Definiendo la función $\hat{g}_t : A_t \times B_t \to \mathbb{R}^{n_t+m_t}$, $\hat{g}_t(x,u) := (-p_t(x), -q_t(u))$, obtenemos equivalentemente

$$X_t \times U_t := \{(x,u) \in A_T \times B_t \; : \; \hat{g}_t(x,u) \geq 0\}.$$

Con estas nuevas conveciones sobre el problema $\mathcal{P}_T(\hat{x}_0)$ vamos a transformarlo en un problema estático. Para ello sea

$$W := A_1 \times \cdots \times A_T \times B_0 \times \cdots \times B_{T-1}.$$

Sus elementos se denotarán como $w := (x_1, \cdots, x_T, u_0, \cdots, u_{T-1})$ (desde que $X_t \subseteq A_t$, $U_t \subseteq B_t$, para no complicar innecesariamente la notación). Notemos que $W \subseteq \mathbb{R}^{(n+m)N}$.

Definamos la función objetivo $F^0 : W \to \mathbb{R}$, y las funciones restricciones $F_t^i : W \to \mathbb{R}$, $G_t^i : W \to \mathbb{R}$ según:

$$F^0(x_1, \cdots, x_T, u_0, \cdots, u_{T-1}) = \textstyle\sum_{k=0}^{T-1} f_k^0(x_k, u_k) + f_T^0(x_T),$$

$$F_t^i(x_1, \cdots, x_T, u_0, \cdots, u_{T-1}) = f_t^i(x_t, u_t) - x_{t+1}^i,$$

$$G_t^i(x_1, \cdots, x_T, u_0, \cdots, u_{T-1}) = g_t^i(x_t, u_t),$$

donde $t = 0, \cdots, T-1,\; ; i = 1, \cdots, n_t$. Con $n_t \in \mathbb{N}$ y $x_0 := \hat{x}_0$.

Con estas definiciones podemos reformular el problema de control óptimo discreto $\mathcal{P}_T(\hat{x}_0)$,

$$\mathcal{P}_T(\hat{x}_0) \quad \equiv \quad \max_{\substack{F_k^i(w)=0, \\ G_t^{i_t}(w) \geq 0.}} F^0(w), \qquad\qquad P_E(Mix)$$

con $k = 0, \cdots, T-1$, $t = 0, \cdots, T-1$ e $i_t = 1, \cdots, n_t$.

Usando este problema equivalente obtenemos el siguiente teorema.

Teorema 4.2.2 (Principio Fuerte de Pontryagin: Programas Mixtos) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ supongamos que su problema equivalente $\mathcal{P}_E(Mix)$ sea mixto en todos sus puntos. Sea $\hat{z} = (\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ una* **solución óptima** *tal que $\hat{x}_t \in \operatorname{int} X_t$, para $t = 1, \cdots, T$.*

Entonces existen $a \in \mathbb{R}$ y vectores $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, $\{\hat{\nu}_s \in \mathbb{R}^{n_s} \; : \; s = 1, \cdots, T\}$ **no todos nulos** *tales que:*

1. *Para $t = 0, \cdots, T$; $i_t = 1, \cdots, n_t$,* $\qquad a \geq 0$ *y* $\hat{\nu}_t^{i_t} \geq 0$,

2. *Para* $t = 0, \cdots, T-1$*;* $i_t = 1, \cdots, n_t$*,*

$$\hat{\nu}_t^{i_t} g_t^{i_t}(\hat{x}_t, \hat{u}_t) = 0.$$

3. *Para* $i_T = 1, \cdots, n_T$*,* $\quad \hat{\nu}_T^{i_T} g_T^{i_T}(\hat{x}_T) = 0.$

4. *Para* $t = 1, \cdots, T-1$*,*

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a, \hat{\psi}_t, \hat{\nu}_t), \qquad (EA - Mix)$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T}[f_T^0(x_T) - \langle \hat{\nu}_T, g_T(x_T) \rangle]$.

5. *Para* $t = 0, \cdots, T-1$*,* $\mathbf{\hat{u}_t}$ *es solución del problema*

$$\max_{u \in U_t} H_t(\hat{x}, u, a, \hat{\psi}_t, \hat{\nu}_t), \qquad (PM)_t(Mix)$$

donde $H_t(x_t, u_t, a, \psi_t, \nu_t) = a f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle - \langle \nu_t, g_t(x_t, u_t) \rangle$ *son los hamiltonianos.*

Prueba. La prueba consiste en aplicar la condición necesaria de optimalidad (Teorema 4.2.1), al problema equivalente $P_E(Mix)$. Desde la optimalidad de $\hat{z}$, tenemos que $\bar{z} := (\hat{x}_1, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ es óptimo del problema $P_E(Mix)$. Por el Teorema 4.2.1 tenemos que existen $a \in \mathbb{R}$ y vectores $\psi_0, \cdots, \psi_{T-1} \in \mathbb{R}^n$, $\{\nu_s \in \mathbb{R}^{n_s} : \ s = 1, \cdots, T\}$ **no todos nulos** tales que:

1. $a \geq 0, \quad \nu_t^{i_t} \leq 0,$ $\qquad$ para $t = 0, \cdots, T$ y $i_t = 1, \cdots, n_t$.

2. $\nu_t^{i_t} \cdot g_t^{i_t}(\hat{x}_t, \hat{u}_t) = 0,$ $\qquad$ para $t = 0, \cdots, T-1$ y $i_t = 1, \cdots, n_t$.

3. Para $i_T = 1, \cdots, n_T,$ $\qquad \nu_T^{i_T} g_T^{i_T}(\hat{x}_T) = 0,$

4. Para todo $x \in X$,

$$\sum_{k=1}^{n} (x^k - \bar{x}^k) \frac{\partial L}{\partial x^k}(\bar{x}, \bar{u}, a, \psi_0, \cdots, \psi_{T-1}, \nu_0, \cdots, \nu_T) \leq 0,$$

donde $\bar{x} := (\hat{x}_1, \cdots, \hat{x}_T)$ y $\bar{u} := (\hat{u}_0, \cdots, \hat{u}_{T-1})$.

5. $(\hat{u}_0, \cdots, \hat{u}_{T-1})$ es solución del problema

$$\max_{(u_0, \cdots, u_{T-1}) \in B} L(\bar{x}, u, a, \psi_0, \cdots, \psi_{T-1}, \nu_0, \cdots, \nu_T),$$

con $B := B_0 \times \cdots \times B_{T-1}$.

En vista de que por hipótesis $\hat{x}_t \in \operatorname{int} X_t$, para $t = 1, \cdots, T$. Entonces de **4.** se deduce

$$\frac{\partial L}{\partial x^k}(\bar{x}, \bar{u}, a, \psi_0, \cdots, \psi_{T-1}, \nu_0, \cdots, \nu_T) = 0. \qquad (\phi)$$

Por su definición podemos expresar el lagrangiano como sigue

$$aF^0(\bar{x}, \bar{u}) + \sum_{t=0}^{T-1} \sum_{k=1}^{n} \psi_t^k F_t^k(\bar{x}, \bar{u}) + \sum_{t=0}^{T} \sum_{k=1}^{n} \nu_t^i G_t^i(\bar{x}, \bar{u}),$$

donde $\psi(t) = (\psi_t^1, \cdots, \psi_t^n)$, $\nu_t = (v_t^1, \cdots, \nu_t^{n_t})$. Calculando su derivada parcial para $k = 1, \cdots, T-1$, y usando (ϕ) tenemos

$$0 = a\frac{\partial f_k}{\partial x^k}(\hat{x}_k, \hat{u}_k) - \sum_{r=1}^{n} \psi_{k-1}^r e^r + \sum_{k=0}^{T} \psi_k^r (\frac{\partial f_k^r}{\partial x^k}(\hat{x}_k, \hat{u}_k)) + \sum_{i_k=1}^{n_k} \nu_k^{i_k} (\frac{\partial g_k^{i_k}}{\partial x^k}(\hat{x}_k, \hat{u}_k)).$$

Defiendo $\hat{\psi}_t := \psi_t, \hat{\nu}_t := -\nu_t$ y $H_t(x_t, u_t, a, \psi_t, \nu_t) = af_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle - \langle \nu_t, g_t(x_t, u_t) \rangle$. Obtenemos para $k = 1, \cdots, T-1$,

$$\hat{\psi}_{k-1} = \nabla_{x_k}(H_k(\hat{x}_k, \hat{u}_k, a, \hat{\psi}_k, \hat{\nu}_k).$$

Similarmente, cuando $k = T$, obtenemos

$$0 = \frac{\partial L}{\partial x^T}(\bar{x}, \bar{u}, a, \psi, \nu) = \frac{\partial}{\partial x^T} f_T^0(\hat{x}_T) - \sum_{r=1}^{n} \psi_{k-1}^r e^r + \sum_{i_k=1}^{n_k} \nu_k^{i_k} (\frac{\partial}{\partial x^k} g_k^{i_k}(\hat{x}_k, \hat{u}_k)).$$

Haciendo $\hat{\psi}_{T-1} := \psi_{T-1}$ y $\hat{\nu}_T := -\nu_T$, se obtiene la condición final

$$\hat{\psi}_{T-1} = \nabla_{x_T}[f_T^0(x_T) - \langle \hat{\nu}_T, g_T(x_T) \rangle].$$

Finalmente en **5.** consideremos elementos de B para cada $u_t \in U_t$,

$$(\hat{u}_0, \cdots, \hat{u}_{t-1}, u_t, \hat{u}_{t+1}, \cdots, \hat{u}_{T-1}) \in B.$$

Calculando el lagrangiano y simplificando en la condición **5.** se obtiene que para toda $u_t \in U_t$, con $t = 0, \cdots, T-1$,

$$R(\hat{u}_t) \geq R(u_t),$$

donde $R(u_t) = af_t(\hat{x}_t, u_t) + \langle \psi_t, f_t(\hat{x}_t, u_t) \rangle + \langle \nu_t, g_t(\hat{x}_t, u_t) \rangle$. Teniendo en cuenta las definiciones de $\hat{\psi}_t$ y $\hat{\nu}_t$, es claro que

$$R(u_t) = af_t^0(\hat{x}_t, u_t) + \langle \hat{\psi}_t, f_t(\hat{x}_t, u_t) \rangle - \langle \hat{\nu}_t, g_t(\hat{x}_t, u_t) \rangle = H_t(\hat{x}_t, u_t, a, \hat{\psi}_t, \hat{\nu}_t).$$

De donde se deduce el ítem **5.** de nuestro teorema. Finalmente los ítems **1.**, **2.** y **3.** se deducen de sus correspondientes. ■

Con la ayuda de este teorema se pueden probar los siguientes corolarios que s e asemejan más al PFP que queremos estudiar.

Corolario 4.2.3 *Para el problema $\mathcal{P}_T(\hat{x}_0)$ satisfaciendo las hipótesis del teorema anterior.*

Supongamos la siguiente hipótesis

- *Para* $t = 0, \cdots, T-1$, *existen funciones* $p_t : A_t \to \mathbb{R}^{m_t}$, $q_t : B_t \to \mathbb{R}^{n_t - m_t}$, *con* $n_t \geq m_t + 1$, *tales que la función* $g_t : A_t \times B_t \to \mathbb{R}^{n_t}$, *satisface* $g_t(x_t, u_t) = (p_t(x_t), q_t(u_t))$.

Sea $\hat{z} = (\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ *una* **solución** *tal que* $\hat{x}_t \in \operatorname{int}(X_t)$, *para* $t = 1, \cdots, T$. *Entonces existen* $a \in \mathbb{R}$ *y unos vectores* $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$, $\{\hat{\nu}_s \in \mathbb{R}^{n_s - m_s} : s = 1, \cdots, T\}$ **no todos nulos** *tales que:*

1. *Para* $t = 0, \cdots, T$, *y* $i_t = 1, \cdots, n_t - m_t$, $\quad a \geq 0, \quad \hat{\nu}_t^{i_t} \geq 0.$

2. *Para* $t = 0, \cdots, T-1$, *y* $i_t = 1, \cdots, n_t - m_t$: $\quad \hat{\nu}_t^{i_t} q_t^{i_t}(\hat{u}_t) = 0.$

3. *Para* $t = 1, \cdots, T-1$,

$$\hat{\psi}_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a, \hat{\psi}_t), \qquad (EA - Mix)$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T} f_T^0(x_T)$.

4. *Para* $t = 0, \cdots, T-1$, $\mathbf{\hat{u}_t}$ *es solución del problema*

$$\max_{u \in U_t} H_t(\hat{x}, u; a, \hat{\psi}_t, \hat{\nu}_t), \qquad (PM)_t(Mix)$$

Donde $H_t(x_t, u_t, a, \psi_t, \nu_t) = a f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle - \langle \nu_t, q_t(u_t) \rangle$ *son los hamiltonianos.*

Prueba. Basta aplicar el teorema anterior (Teorema 4.2.2) bajo las condiciones de este corolario. En particular deducímos desde $\hat{x}_t \in \operatorname{int} X_t$, que $p_t(x_t) > 0$ y por ello desde la complementariedad los primeros multiplicadores son nulos $\hat{\nu}_t^{i_t} = 0$, con $t = 1, \cdots, T$, e $i_t = 1, \cdots, m_t$, de este hecho los hamiltonianos se reducirán y sólo importarán $\hat{\nu}_t^{i_t}$, cuando $i_t = m_t + 1, \cdots, n_t$, redefiniendo estos multiplicadores

$$\hat{\nu}_t^{i_t} := \hat{\nu}_t^{i_t - m_t}$$

se deduce todo el teorema. ■

Corolario 4.2.4 (Interioridad) *Para el problema $\mathcal{P}_T(\hat{x}_0)$ con las condiciones del corolario anterior. Sea* $\hat{z} = (\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ *una* **solución óptima** *tal que para* $t = 0, \cdots, T-1$,

$$\hat{u}_t \in \operatorname{int}(U_t), \qquad y \qquad \hat{x}_{t+1} \in \operatorname{int}(X_{t+1}).$$

Entonces existen $a \in \mathbb{R}$ *y vectores* $\hat{\psi}_0, \cdots, \hat{\psi}_{T-1} \in \mathbb{R}^n$ **no todos nulos** *tales que:*

1. $a \geq 0$,

2. *Para* $t = 1, \cdots, T-1$,

$$\hat{\psi}_{t-1} = \nabla_{x_t}(H_t(\hat{x}_t, \hat{u}_t, a, \hat{\psi}_t), \qquad (EA - Mix)$$

con condición final, $\hat{\psi}_{T-1} = \nabla_{x_T} f_T^0(x_T)$.

4. *Para* $t = 0, \cdots, T-1$, $\mathbf{\hat{u}_t}$ *es solución del problema*

$$\max_{u \in U_t} H_t(\hat{x}, u, a, \hat{\psi}_t), \qquad (PM)_t(Mix)$$

donde $H_t(x_t, u_t, a, \psi_t) = a f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$ *son los hamiltonianos.*

Prueba. Nuevamente la prueba se obtiene al ver que $\hat{u}_t \in \operatorname{int}(U_t)$, entonces los multiplicadores $\hat{\nu}_t$ son nulos. Entonces del corolario anterior se deduce el actual. ■

Observación. Vemos que estos principios se asemejan a los resultados obtenidos en el Capítulo 3, utilizando la aproximación de Karush-Kuhn-Tucker. Probemos en último lugar un corolario que hace más evidente este hecho.

Corolario 4.2.5 (PFP Mixto Convexo) *Para el problema* $\mathcal{P}_T(\hat{x}_0)$ *supongamos las siguientes hipótesis*

- *Para* $t = 0, \cdots, T-1$,

$$f_t \text{ es una función afín en la variable "}u\text{".} \qquad (Afin - u)$$

- *Para* $t = 0, \cdots, T-1$,

$$f_t^0 \text{ es una función convexa en la variable "}u\text{".} \qquad (Cvxa - u)$$

- *Para los conjuntos* $U_t = \{u \in \mathbb{R}^m : q_t(u) \leq 0\}$, *las funciones* q_t *son convexas.*

Entonces el problema $\mathcal{P}_T(\hat{x}_0)$ *es un Problema Mixto en todo punto.*

En consecuencia si $\hat{z} = (\hat{x}_0, \cdots, \hat{x}_T, \hat{u}_0, \cdots, \hat{u}_{T-1})$ *es una* **solución** *tal que* $\hat{x}_t \in \operatorname{int}(X_t)$, *para* $t = 1, \cdots, T$, *entonces se satisface las mismas conclusiones que el Corolario 4.2.3.*

Prueba. Desde las hipótesis $(Afin - u)$, $(Cvxa - u)$ y la última, para el problema estático equivalente $\mathcal{P}_E(Mix)$, se satisface que el conjunto $A(x)$ es convexo por lo que del Corolario 4.2.1 el problema es mixto en todo punto, luego podemos usar el Corolario 4.2.3. ■

Observación. Notemos que estos resultados no llegan a tener la generalidad de Teorema 4.1.1 (PFP-Afín), en el sentido de que para tener el PFP propiamente hablando (Proposición 3.1.2), requiere hipótesis de interioridad de los controles o los estados ($\hat{x}_t \in \operatorname{int}(X_t), \hat{u}_t \in \operatorname{int}(U_t)$), las cuales no son exigidas en el Teorema 4.1.1. Por todo esto ¿Son independientes ambas aproximaciones? la respuesta es sí, veamoslo en la siguiente subsección.

4.2.3 Independencia de PFP Afín y PFP Mixto

Para concluir con esta sección daremos un ejemplo de un problema que es afín en la variable "x" (i.e. las funciones objetivo y restricciones), y que no es un programa mixto. Por lo cual se hace evidente que los PFP según afinidad de "x" (Teorema 4.1.1) y según programas mixtos (Teorema 4.2.2) son independientes entre sí, lo que resultaba ser intuitivo, pues éste último es más restringido que el primero.

Consideremos el siguiente problema

$$\begin{array}{cc} \max & f^0(x,u) \\ s.a & f^1(x,u) \geq 0, \end{array}$$

donde $X = \mathbb{R}$, $U = [-2,2]$, y las funciones $f^0, f^1 : X \times U \to \mathbb{R}$, están definidas según

$$f^0(x,u) := \begin{cases} 2u, & \text{si } u \in [-2,1/2], \\ 2-2u, & \text{si } u \in [1/2,1], \\ 24(u-1)x + 24(\frac{u-1}{\sqrt{u}})^2, & \text{si } u \in [1,2], \end{cases}$$

$$f^1(x,u) := 1 - ux - u.$$

De sus definiciones es claro que las funciones f^0 y f^1 son afines respecto a la variable "x".

Probemos que el punto $(\hat{x}, \hat{u}) := (0, 1/2)$ **es una solución óptima** del problema. En efecto $f^0(0,1/2) = 1$, de su propia definición para todo (x,u) tal que $u \in [-2,1]$ se tiene que $f^0(x,u) \leq 1$. Por su parte, cuando $u \in [1,2]$,

$$f^0(x,u) = 24(u-1)\left[x + \frac{u-1}{u}\right] = 24(u-1)\left[\frac{ux+u-1}{u}\right] \leq 0,$$

pues $1 - ux - u = f^1(x,u) \geq 0$.

Ahora probemos que este problema no es mixto en el punto $(\hat{x}, \hat{u}) := (0, 1/2)$, para ello usaremos la condición necesaria dada en el Lema 4.2.3.

Calculando el conjunto $\mathcal{U}(\hat{x}) := \{u \in [-2,2] : |f^1(\hat{x},u)| \leq |f^1(\hat{x},\hat{u})|\}$. Obtenemos el intervalo $\mathcal{U}(0) = [\frac{1}{2}, \frac{3}{2}]$.

Del Lema 4.2.3, si el problema fuese mixto en $(0, 1/2)$, entonces para todo $u \in \mathcal{U}(0) = [\frac{1}{2}, \frac{3}{2}]$ debe cumplirse

$$1 = f^0(0, 1/2) \geq f^0(0, u),$$

lo cual es falso pues para $u = 3/2$ se tiene $f^0(0, 3/2) = 2$. Por lo tanto el problema no es mixto en $(0, 1/2)$, así que no es mixto en todo punto. ■

4.3 Condición Suficiente de KKT

La tercera aproximación para obtener el PFP, la hemos realizado en los Corolarios 3.4.2 y 3.4.3 del Capítulo 3 en los cuales se obtienen al dar condiciones al problema $P_T(\hat{x}_0)$, **que la condición necesaria** del Teorema 3.4.1, esta es:

Para $t = 0, \cdots, T-1$, $\hat{u}_t$ es **punto crítico** del problema

$$\max_{u_t \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t, \hat{\nu}_t), \qquad (PMdPK)_t$$

sea también una condición suficiente, y $\hat{u}_t$ sea de hecho una solución del problema.

Finalizemos esta sección resaltando que el PFP planteado en el capítulo 3 (Proposición 3.1.2), es satisfecho propiamente sólo cuando las funciones satisfacen las hipótesis de afinidad respecto a la variable "x" (ver Teoremas 4.1.1 y 4.1.2), mientras que las otras 2 aproximaciones (Programas Mixtos y KKT) requieren hipótesis de interioridad (de los estados x_t o controles u_t). Por ello consideramos que la investigación realizada en la sección 1 es relevante en este sentido, más si tenemos en cuenta la relación que se encuentra entre la Programación Dinámica y el PFP en la demostración de estos teoremas, la cual nos hace ver porque a lo largo del capítulo 3 hemos logrado demostrar los Principios Débiles. Es decir, los Principios Débiles son ciertos porque en lo profundo, la relación que hay entre la Programación Dinámica y el PFP, se siguen manteniendo pero a un nivel más débil.

Capítulo 5

Horizonte Infinito

En este capítulo extenderemos los resultados obtenidos para el problema de horizonte finito $P_T(\hat{x}_0)$, a los problemas de horizonte infinito $P(\hat{x}_0)$ definidos en el Capítulo 1.

En la última parte de este capítulo logramos demostrar un resultado a nuestro juicio importante el Principio Fuerte de Pontryagin en Horizonte Infinito.

Con este objetivo, recordemos primeramente la definición exacta de estos problemas dada en dicho capítulo.

$$\begin{array}{rc} \max & B((x_0, x_1, \cdots), (u_0, \cdots)) := \sum_{t=0}^{\infty} f_t^0(x_t, u_t) \\ \text{s.a} & \text{para todo} \;\; t \geq 0, \; x_t \in X_t, \, u_t \in U_t, \\ & x_{t+1} = f_t(x_t, u_t), \; x_0 = \hat{x}_0 \in X_0. \end{array} \tag{$\mathcal{P}(\hat{x}_0)$}$$

donde, similar al caso de horizonte finito, para todo $t \geq 0$,

- $X_t \subseteq \mathbb{R}^n$, $U_t \subseteq \mathbb{R}^m$,
- $f_t^0 : X_t \times U_t \to \mathbb{R}, \; f_t : X_t \times U_t \to \mathbb{R}^n$.

El conjunto de los $((x_0, x_1, \cdots), (u_0, \cdots))$ satisfaciendo el sistema en el problema $\mathcal{P}(\hat{x}_0)$ es llamado conjunto admisible y es denotado por $\operatorname{Adm}(\hat{x}_0)$,

$$\operatorname{Adm}(\hat{x}_0) := \left\{ ((x_0, x_1, \cdots), (u_0, \cdots)) \;\middle|\; \begin{array}{l} \text{para todo} \;\; t \geq 0, \, x_t \in X_t, \;\, u_t \in U_t, \\ x_{t+1} = f_t(x_t, u_t) \;\text{ y }\; x_0 = \hat{x}_0 \in X_0. \end{array} \right\}.$$

Junto con este problema también podemos definir 2 problemas adicionales según usemos como criterio de optimalidad ($\mathcal{L}, \mathcal{I}$ y $\mathcal{L}, \mathcal{II}$) (ver Capítulo 1), los denotaremos $\mathcal{P}_{\mathcal{I}}(\hat{x}_0)$ y $\mathcal{P}_{\mathcal{II}}(\hat{x}_0)$. En vista de que la metología de solución es la misma para estos problemas, no haremos mucha distinción entre los mismos.

5.1 Programación Dinámica

Para simplificar los resultados de esta sección, al igual en el Capítulo 2 tenemos que para un elemento $x_t \in X_t$ y una sucesión de decisiones $(u_{t+1}, u_{t+2}, \cdots)$, se determina un único proceso $(x_t, x_{t+1}, \cdots)$, por ello la función beneficio a partir del tiempo "t" (B^t) puede denotarse simplificadamente

$$\sum_{s\geq t} f^0(x_s, u_s) = B^t(x_t, x_{t+1}, \cdots; u_t, u_{t+1}, \cdots) := B^t(x_t, u_t, u_{t+1}, \cdots).$$

En consecuencia podemos definir las secuencias de decisiones admisibles para un elemento $\hat{x}_t \in X_t$ como sigue

$$\mathrm{Adm}\,(\hat{x}_t) := \left\{ (u_t, u_{t+1}, \cdots) \;\middle|\; \begin{array}{l} \text{para } \; s \geq t: \; u_s \in U_s, \; \hat{x}_t = x_t, \\ x_{s+1} = f_s(x_s, u_s) \in X_{s+1}, \;\text{ y} \\ \sum_{s\geq t} f_s^0(x_s, u_s) \quad \text{converge} \end{array} \right\}.$$

Usando este conjunto podemos definir el subproblema,

$$\max_{\mathrm{Adm}\,(\hat{x}_t)} B^t(\hat{x}_t, u_t, u_{t+1}, \cdots), \qquad \mathcal{P}^t(\hat{x}_t)$$

al cual asociamos su función valor,

$$\begin{array}{rccl} V_t: & X_t & \to & \mathbb{R} \\ & x_t & \mapsto & \sup_{\mathrm{Adm}\,(x_t)} B^t(x_t, u_t, u_{t+1}, \cdots). \end{array}$$

Estas funciones cumplen las Ecuaciones de Bellman.

Proposición 5.1.1 (Ecuaciones de Bellman) *Para cada* $t \geq 0$ *y* $x_t \in X_t$,

$$V_t(x_t) = \sup_{u_t \in U_t} \left\{ f_t^0(x_t, u_t) + V_{t+1}(f_t(x_t, u_t)) \right\}. \qquad (EBi_t)$$

Prueba. Tomemos $t \geq 0$ y $x_t \in X_t$ fijos arbitrarios. Demostremos la igualdad en dos partes. Primero dado u_t fijo arbitrario, por la definición de V_{t+1} como un supremo, dado $\epsilon > 0$, existe una secuencia de decisiones $(\hat{u}_{t+1}, \hat{u}_{t+2}, \cdots)$ tal que

$$V_{t+1}(f_t(x_t, u_t)) - \epsilon < B^{t+1}(f_t(x_t, u_t); \hat{u}_{t+1}, \hat{u}_{t+2}, \cdots),$$

sumando a ambos lados $f_t^0(x_t, u_t)$, y usando la definición de B_t tenemos

$$f_t^0(x_t, u_t) + V_{t+1}(f_t(x_t, u_t)) - \epsilon < B^t(x_t; \hat{u}_t, \hat{u}_{t+1}, \hat{u}_{t+2}, \cdots) \leq V_t(x_t).$$

Como $\epsilon > 0$ puede ser tomado arbitrariamente pequeño, deducimos que para cualquier $u_t \in U_t$,

$$f_t^0(x_t, u_t) + V_{t+1}(f_t(x_t, u_t)) \leq V_t(x_t),$$

por lo cual

$$\sup_{u_t\in U_t}\left\{f_t^0(x_t,u_t)+V_{t+1}(f_t(x_t,u_t))\right\}\leq V_t(x_t).$$

Por otra parte tomando nuevamente $\epsilon>0$ arbitrario, desde que V_t es supremo existe una secuencia de decisiones $(\hat{u}_t,\hat{u}_{t+1},\cdots)$ tal que

$$B^t(x_t;\hat{u}_t,\hat{u}_{t+1},\cdots)\geq V_t(x_t)-\epsilon.$$

Por la definición de B^{t+1} es claro que

$$\begin{aligned}B^t(x_t;\hat{u}_t,\hat{u}_{t+1},\cdots) &= f_t^0(x_t,\hat{u}_t)+B^{t+1}(f_t(x_t,\hat{u}_t);\hat{u}_{t+1},\cdots)\\ &\leq \sup_{u_t\in U_t}\left\{f_t^0(x_t,u_t)+V_{t+1}(f_t(x_t,u_t))\right\}.\end{aligned}$$

Desde que $\epsilon>0$ es abitrario deducimos

$$V_t(x_t)\geq\sup_{u_t\in U_t}\left\{f_t^0(x_t,u_t)+V_{t+1}(f_t(x_t,u_t))\right\},$$

por lo que juntando ambas desigualdades hemos probado la igualdad. ■

Caracterizemos ahora, las soluciones óptimas usando estas ecuaciones, teniendo en cuenta que también en este contexto se cumple el principio de optimalidad de Bellman ("Dada una secuencia óptima de decisiones, toda subsecuencia de ella es, a su vez, óptima"), por ello deducimos el siguiente resultado.

Proposición 5.1.2 (Principio de Optimalidad de Bellman) *Para el problema $\mathcal{P}(\hat{x}_0)$, sea $(\hat{u}_1,\hat{u}_2,\cdots)$ una solución óptima, con su proceso asociado $(\hat{x}_0,\hat{x}_1,\cdots)$. Entonces para cada $k\geq 0$, la sucesión $(\hat{u}_k,\hat{u}_{k+1},\cdots)$ es solución del problema $\mathcal{P}^k(\hat{x}_k)$.*

El mismo resultado se tiene para los problemas $\mathcal{P}_{\mathcal{I}}(\hat{x}_0)$ y $\mathcal{P}_{\mathcal{I}\mathcal{I}}(\hat{x}_0)$.

Prueba. Probemos el resultado por inducción. Cuando $k=0$ es claro que $(\hat{u}_0,\hat{u}_2,\cdots)$ es solución del problema $\mathcal{P}(\hat{x}_0)\equiv\mathcal{P}^0(\hat{x}_0)$.

Supongamos demostrado el resultado hasta un valor $k\geq 1$, (i.e. $(\hat{u}_k,\hat{u}_{k+1},\cdots)$ es solución del problema $\mathcal{P}^k(\hat{x}_k)$). Consideremos el problema $\mathcal{P}^{k+1}(\hat{x}_{k+1})$. Si $(\hat{u}_{k+1},\hat{u}_{k+2},\cdots)$ no fuese óptima, tendríamos una sucesión $(\bar{u}_{k+1},\bar{u}_{k+2},\cdots)$ tal que

$$B^k(\hat{x}_{k+1};\bar{u}_{k+1},\bar{u}_{k+2},\cdots)>B^k(\hat{x}_{k+1};\hat{u}_{k+1},\hat{u}_{k+2},\cdots).$$

Sumando $f^0(\hat{x}_k,\hat{u}_k)$ a ambos lados tenemos

$$B^k(\hat{x}_k;\hat{u}_k,\bar{u}_{k+1},\bar{u}_{k+2},\cdots)>B^k(\hat{x}_k;\hat{u}_k,\hat{u}_{k+1},\hat{u}_{k+2},\cdots),$$

lo cual contradice la optimalidad de la sucesión $(\hat{u}_k,\hat{u}_{k+1},\cdots)$ para el problema $\mathcal{P}^k(\hat{x}_k)$. Por lo tanto tiene que ser óptima.

Para los problemas $\mathcal{P}_{\mathcal{I}}(\hat{x}_0)$ y $\mathcal{P}_{\mathcal{II}}(\hat{x}_0)$, la argumentación es similar, teniendo en cuenta que la expresión

$$f_t^0(\hat{x}_t, \hat{u}_t) - f_t^0(x_t, u_t) + \liminf_{T\to\infty} \left(\sum_{s=t+1}^{T} f_s^0(\hat{x}_s, \hat{u}_s) - \sum_{s=t+1}^{T} f_s^0(x_s, u_s) \right),$$

es igual a

$$\liminf_{T\to\infty} \left(\sum_{s=t}^{T} f_s^0(\hat{x}_s, \hat{u}_s) - \sum_{s=t}^{T} f_s^0(x_s, u_s) \right).$$

Por ello podemos usar la inducción. Similar aproximación se tiene para el lim sup, en el problema $\mathcal{P}_{\mathcal{II}}(\hat{x}_0)$. ■

Con estos conceptos obtenemos la siguiente caracterización de las soluciones óptimas.

Proposición 5.1.3 (Caracterización de la Solución Óptima) *La sucesión* $(\hat{u}_0, \hat{u}_1, \cdots)$, *con su proceso asociado* $(\hat{x}_0, \hat{x}_1, \cdots)$ *es una solución óptima del problema* $\mathcal{P}(\hat{x}_0)$ *si y sólo si para cada* $t \geq 0$,

$$\begin{aligned} V_t(\hat{x}_t) &= f_t^0(\hat{x}_t, \hat{u}_t) + V_{t+1}(f_t(\hat{x}_t, \hat{u}_t)) \\ &= \sup_{u_t \in U_t} \left\{ f_t^0(\hat{x}_t, u_t) + V_{t+1}(f_t(\hat{x}_t, u_t)) \right\}, \end{aligned}$$

y además la condición de transversalidad $\lim_{T\to\infty} V_T(\hat{x}_T) = 0$.

Prueba. ($\Rightarrow$) Tomemos $t \geq 0$ fijo y arbitrario. Del Principio de optimalidad de Bellman, tenemos que $(\hat{u}_{t+1}, \hat{u}_{t+2}, \cdots)$ es solución del problema $\mathcal{P}^{t+1}(\hat{x}_{t+1})$, por ello

$$V_{t+1}(f_t(\hat{x}_t, \hat{u}_t)) = V_{t+1}(\hat{x}_{t+1}) = B^{t+1}(\hat{x}_{t+1}, \hat{u}_{t+1}, \hat{u}_{t+2}, \cdots). \tag{1}$$

Similarmente $(\hat{u}_t, \hat{u}_{t+1}, \cdots)$ es solución del problema $\mathcal{P}^t(\hat{x}_t)$, esto es

$$V_t(\hat{x}_t) = B^t(\hat{x}_t, \hat{u}_{t+1}, \hat{u}_{t+2}, \cdots) = f_t^0(\hat{x}_t, \hat{u}_t) + B^{t+1}(\hat{x}_{t+1}, \hat{u}_{t+1}, \hat{u}_{t+2}, \cdots)$$

Usando **(1)** obtenemos

$$V_t(\hat{x}_t) = f_{t+1}^0(\hat{x}_t, \hat{u}_{t+1}) + V_{t+1}(f_{t+1}(\hat{x}_t, \hat{u}_{t+1})).$$

Para deducir la primera implicación, sólo nos falta probar que se cumple la condición de Transversalidad.

Desde las igualdades $V_t(\hat{x}_t) = f_t^0(\hat{x}_t, \hat{u}_t) + V_{t+1}(f_t(\hat{x}_t, \hat{u}_t))$ es fácil obtener por inducción que para todo $T \geq 1$ se tiene

$$V_0(\hat{x}_0) - \sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t) = V_T(\hat{x}_T). \tag{2}$$

Desde la optimalidad de las decisiones $(\hat{u}_0, \hat{u}_1, \cdots)$, de la definición de V_0, y de la identidad (2) se tiene

$$0 = V_0(\hat{x}_0) - \sum_{t=1}^{\infty} f_t^0(\hat{x}_t, \hat{u}_t) = \lim_{T\to\infty} \left(V_0(\hat{x}_0) - \sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t) \right) = \lim_{T\to\infty} V_T(\hat{x}_T).$$

($\Leftarrow$) Recíprocamente si un camino cumpliese para cada $t \geq 0$,

$$V_t(\hat{x}_t) = f_t^0(\hat{x}_t, \hat{u}_t) + V_{t+1}(f_t(\hat{x}_t, \hat{u}_t)),$$

inductivamente tenemos para todo $T \geq 0$,

$$V_0(\hat{x}_0) = \sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t) + V_T(\hat{x}_T). \tag{2}$$

De otro lado $(\hat{u}_0, \hat{u}_1, \cdots)$ es admisible entonces $\sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t)$ converge. Teniendo en cuenta esto, la condición de transversalidad y la igualdad (2) deducímos,

$$\begin{array}{rcl} \lim_{T\to\infty} [V_0(\hat{x}_0) - V_T(\hat{x}_T)] & = & \lim_{T\to\infty} \left[\sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t)\right], \\ V_0(\hat{x}_0) - \lim_{T\to\infty} V_T(\hat{x}_T) & = & \sum_{t=1}^{\infty} f_t^0(\hat{x}_t, \hat{u}_t). \end{array}$$

Por tanto,

$$V_0(\hat{x}_0) = \sum_{t=0}^{\infty} f_t^0(\hat{x}_t, \hat{u}_t),$$

de donde por la definición de V_0, la sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$ es solución óptima. ■

Luego de haber hecho todo lo anterior podríamos pensar que al igual que en el caso finito también podemos obtener un método que nos permite resolver al menos teóricamente el problema $\mathcal{P}(\hat{x}_0)$. Pero a diferencia del caso finito aquí sucede un hecho muy especial y es que los subproblemas siguen siendo infinitos por lo que continuan manteniendo su complejidad. No así en el caso finito, en el que los subproblemas podían ordenarse desde simples a complejos. Y se procedía a resolverlos en orden de creciente complejidad.

Sin embargo es posible usar los problemas finitos para aproximar soluciones para el problema de horizonte infinito. Para ello es preciso tener en cuenta la condición de transversalidad ($\lim_{T\to\infty} V_T(\hat{x}_T) = 0$). En esta dirección propongamos la siguiente definición para el problema $\mathcal{P}(\hat{x}_0)$. Diremos que una sucesión de decisiones $(\hat{u}_0, \hat{u}_1, \cdots, \hat{u}_{T-1})$ es una solución **ϵ-óptima** de $\mathcal{P}(\hat{x}_0)$, si

$$V_0(\hat{x}_0) - \epsilon < \sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t).$$

Proposición 5.1.4 *Para el problema $\mathcal{P}(\hat{x}_0)$ supongamos que $V_0(\hat{x}_0) < \infty$. Entonces para todo $\epsilon > 0$ existe $T \in \mathbb{N}$ suficientemente grande tal que la solución del problema finito $\mathcal{P}_T(\hat{x}_0)$ es una solución* ϵ**-óptima** *de $\mathcal{P}(\hat{x}_0)$.*

Prueba. Dado $\epsilon > 0$, desde la definición de $V_0(\hat{x}_0)$ como supremo existe una sucesión admisible $(\bar{u}_0, \bar{u}_1, \cdots,)$, con su camino asociado $(\hat{x}_0, \bar{x}_1, \cdots)$ tal que

$$V_0(\hat{x}_0) - \epsilon/2 < B(\hat{x}_0; \bar{u}_0, \bar{u}_1, \cdots). \tag{1}$$

Desde la convergencia de la sumatoria $\sum_{t=0}^{T-1} f^0(\hat{x}_t, \hat{u}_t)$, existe un T suficientemente grande tal que

$$\left| \sum_{t=T}^{\infty} f^0(\bar{x}_t, \bar{u}_t) \right| < \epsilon/2. \tag{2}$$

Por otra parte la sucesión $(\bar{u}_0, \bar{u}_1, \cdots, \bar{u}_{T-1})$ es admisible en el problema $\mathcal{P}_T(\hat{x}_0)$, pues lo es para $\mathcal{P}(\hat{x}_0)$.

Tomemos una solución óptima $(\hat{u}_0, \hat{u}_1, \cdots, \hat{u}_{T-1})$ del problema $\mathcal{P}_T(\hat{x}_0)$, (existe desde que las sumatorias $\sum_{t=0}^{T-1} f^0(\hat{x}_t, \hat{u}_t)$ están acotadas por $V_0(\hat{x}_0) < \infty$) entonces

$$B_T(\hat{x}_0, \bar{u}_0, \bar{u}_1, \cdots, \bar{x}_{T-1}) < B_T(\hat{x}_0, \hat{u}_0, \hat{u}_1, \cdots, \hat{x}_{T-1}).$$

De las desigualdades (1), (2) y de ésta última deducimos la siguiente

$$V_0(\hat{x}_0) - \epsilon < B_T(\hat{x}_0; \hat{u}_0, \hat{u}_1, \cdots, \hat{x}_{T-1}),$$

por lo que $(\hat{u}_0, \hat{u}_1, \cdots, \hat{x}_{T-1})$ es ϵ-óptima. ∎

El teorema anterior nos permite rescatar los resultados del capítulo 2 en el estudio de los problemas de horizonte infinito.

Sin embargo dentro de la Programación Dinámica es posible realizar otra aproximación al estudio de las soluciones del problema $\mathcal{P}(\hat{x}_0)$, la cual se obtiene al observar que de hecho todas las funciones valor son iguales. Tenemos el siguiente lema

Lema 5.1.1 (Función Valor) *Para el problema $\mathcal{P}(\hat{x}_0)$ supongamos que las funciones de pago sean una sola, es decir $f_t^0 = f^0$, para cada $t \geq 0$. Sean V_t las funciones valor de los subproblemas $\mathcal{P}^t(\hat{x}_0)$ entonces todas ellas son iguales. Por ello asociada a $\mathcal{P}(\hat{x}_0)$ se define una única función valor V.*

Prueba. La prueba es evidente a partir de la hipótesis pues dados t_1 y t_2

$$B_{t_1}(x_{t_1}, u_{t_1}, u_{t_1+1}, \cdots) = \sum_{s \geq t_1} f^0(x_s, u_s) = \sum_{s \geq t_2} f^0(x_s, u_s) = B_{t_2}(x_{t_2}, u_{t_2}, \cdots),$$

de la cual se deduce la igualdad $V_{t_1} = V_{t_2}$. ∎

Una vez obtenido el lema anterior es posible estudiar ciertos resultados que caracterizan el comportamiento de esta función valor. Dichos resultados e investigaciones se encuentran en el libro celebre de Richard Bellman "Programación Dinámica (ver [2])". No desarrollaremos estos resultados en esta Tesis. Pasaremos en cambio al estudio de otro método para abordar el problema de control discreto, el cual consiste en extender el principio débil de Pontryagin discreto para los problemas de horizonte infinito.

5.2 Principio de Pontryagin Lagrange Infinito

En esta sección, bajo ciertas hipótesis para el problema $\mathcal{P}(\hat{x}_0)$, probaremos una condición necesaria de optimalidad, que por su forma, generaliza el Principio Débil de Pontryagin Lagrange (PdPL) (Teorema 3.2.1), para problemas en horizonte infinito.

Con este objetivo primeramente "reduciremos" un problema infinito a una familia de problemas finitos, a partir de los cuales trataremos de generar una condición necesaria de optimalidad.

Lema 5.2.1 (Lema de Reducción a Problemas Finitos) *Para $\mathcal{P}(\hat{x}_0)$, sea la sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$, con su proceso asociado $(\hat{x}_0, \hat{x}_1, \cdots)$, una solución óptima, entonces para todo $T \geq 1$, $(\hat{u}_0, \hat{u}_1, \cdots, \hat{u}_{T-1})$, es una solución del problema de extremo derecho fijo,*

$$\begin{array}{ll} \max & B_T(\hat{x}_0, u_0, \cdots, u_{T-1}) := \sum_{t=0}^{T-1} f_t^0(x_t, u_t) \\ s.a. & \textit{para todo } t = 0, \cdots, T-1,\ u_t \in U_t, \\ & x_{t+1} \in X_{t+1},\ \ x_{t+1} = f_t(x_t, u_t), \\ & x_0 = \hat{x}_0 \in X_0 \ \ y \ \ x_T = \hat{x}_T. \end{array} \qquad (\mathcal{P}_T(\hat{x}_0, \hat{x}_T))$$

Un resultado igual se tiene para los problemas $\mathcal{P}_{\mathcal{I}}(\hat{x}_0)$ y $\mathcal{P}_{\mathcal{II}}(\hat{x}_0)$.

Prueba. Tomemos $T \geq 1$ fijo y arbitrario, por contradicción supongamos que $(\hat{u}_0, \hat{u}_1, \cdots, \hat{u}_{T-1})$, no sea solución entonces existe $(v_0, \cdots, v_{T-1})$ admisible para $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$ con su proceso asociado $(\hat{x}_0, y_1, \cdots, y_{T-1}, \hat{x}_T)$, tales que

$$B_T(\hat{x}_0, v_0, \cdots, v_{T-1}) > B_T(\hat{x}_0, \hat{u}_0, \cdots, \hat{u}_{T-1}).$$

Desde que $\hat{x}_T$ es extremo fijo, la sucesión $(v_0, \cdots, v_{T-1}, \hat{u}_T, \hat{u}_{T+1}, \cdots) \in \mathrm{Adm}\,(\hat{x}_0)$, por lo que tendríamos

$$B(v_0, \cdots, v_{T-1}, \hat{u}_T, \hat{u}_{T+1}, \cdots) - B(\hat{x}_0, \hat{u}_0, \cdots) = \sum_{t=0}^{T-1} f_t^0(y_t, v_t) - \sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t) > 0,$$

lo cual contradice la optimalidad de $(\hat{u}_0, \hat{u}_1, \cdots)$, por lo que el resultado queda demostrado.

Por otra parte si el problema fuese $\mathcal{P}_{\mathcal{I}}(\hat{x}_0)$ entonces es claro que

$$\liminf_{S\to\infty}\left(\sum_{t=0}^{S} f_t^0(y_t, v_t) - \sum_{t=0}^{S} f_t^0(\hat{x}_t, \hat{u}_t)\right) = \sum_{t=0}^{T-1} f_t^0(y_t, v_t) - \sum_{t=0}^{T-1} f_t^0(\hat{x}_t, \hat{u}_t) > 0,$$

por lo tanto nuevamente obtendríamos una contradicción. Finalmente se usa la misma idea para el problema $\mathcal{P}_{\mathcal{II}}(\hat{x}_0)$. ■

Usando el lema anterior podemos demostrar una condición necesaria de optimalidad para las sucesiones que satisfagan las siguientes hipótesis:

- Para todo $t \geq 0$,
$$u_t \in \operatorname{int} \mathrm{U_t}, \quad \mathrm{x_{t+1}} \in \operatorname{int} \mathrm{X_{t+1}}. \qquad (I)$$

- Existe una sucesión $(t_j)_{j\in\mathbb{N}}$ tal que,
$$D_{u_{t_j}} f_{t_j}(x_{t_j}, u_{t_j}), \quad \text{es sobreyectiva.} \qquad (Sur)$$

Además usaremos las hipótesis $(H1)$ y $(H2)$ del Capítulo 3 para todo $t \geq 1$.

Teorema 5.2.1 (Principio de Pontryagin Lagrange Discreto Infinito) *Para el problema $\mathcal{P}(\hat{x}_0)$ satisfaciendo $(H1)$, $(H2)$ del Capítulo 3 para todo $t \geq 1$. Sea una sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$, con su proceso asociado $(\hat{x}_0, \hat{x}_1, \cdots)$, una solución óptima que satisface las hipótesis (I) y (Sur).*

Entonces existen vectores $\psi_0, \psi_1, \psi_2, \cdots \in \mathbb{R}^n$ tales que:

1. *Para $t \geq 1$,*
$$\psi_{t-1} = D_{x_t} H_t(\hat{x}_t, \hat{u}_t, \psi_t). \qquad (EA)_\infty$$

2. *Para $t \geq 0$,*
$$D_{u_t} H_t(\hat{x}_t, \hat{u}_t, \psi_t) = 0. \qquad (PdPL)_\infty$$

Prueba. Desde el lema de reducción a problemas finitos (Lema 5.2.1) tenemos para cada $T \geq 1$, $(\hat{u}_0, \hat{u}_1, \cdots \hat{u}_T)$ es solución de $\mathcal{P}(T, \hat{x}_0, \hat{x}_T)$.

Para cada uno de estos problemas apliquemos el Corolario 3.2.1 del Teorema 3.5.1 (PdPB) del Capítulo 3 para problemas de extremos fijo. Tendremos que existen asociadas a cada uno de ellos los vectores $\psi_1^T, \psi_2^T, \cdots, \psi_T^T \in \mathbb{R}^n$ tales que:

- Para $t = 1, \cdots, T-1$,
$$\psi_{t-1}^T = \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) + \psi_t^T \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t). \qquad (1)_T$$

- Para $t = 0, \cdots, T-1$,

$$\nabla_{u_t} f_t^0(\hat{x}_t, \hat{u}_t) + \psi_t^T \circ D_{u_t} f_t(\hat{x}_t, \hat{u}_t) = 0. \tag{$2)_T$}$$

Tomemos "t" fijo, entonces desde $(2)_T$, para todo $T \geq t$,

$$\psi_t^T \circ D_{u_t} f_t(\hat{x}_t, \hat{u}_t) = -\nabla_{u_t} f_t^0(\hat{x}_t, \hat{u}_t). \tag{3}$$

Para cada $t \geq 0$, a partir del conjunto de vectores $\{\psi_t^T : T \geq t\}$, vamos a definir los vectores "ψ_t". Demostremos en primer lugar que si $t \in S$, (S definido en la hipótesis (Sur)) entonces para todo $T \geq t$, tenemos $\psi_t^T = \psi_t^t$.

Sea $t \in S$, consideremos $T_1, T_2 \geq t$. De la hipótesis (Sur), $D_{u_t} f_t(x_t, u_t)$ es sobreyectivo, por ello su transpuesta $(D_{u_t} f_t(x_t, u_t))^t$ es inyectiva.

Considerando esta inyectividad en la identidad (3) tenemos para $T_1, T_2 \geq t$

$$(D_{u_t} f_t(x_t, u_t))^t \circ \psi_t^{T_1} = (-\nabla_{u_t} f_t^0(\hat{x}_t, \hat{u}_t))^t = (D_{u_t} f_t(x_t, u_t))^t \circ \psi_t^{T_2},$$

en consecuencia $\psi_t^{T_1} = \psi_t^{T_2}$.

Usando este resultado definamos para cada $t \geq 0$, los vectores

$$\psi_t := \psi_t^{t_0},$$

donde $t_0 := \min\{s \in S : s \geq t\}$.

Demostremos que con estos vectores se satisfacen las implicaciones **1.** y **2.** del Teorema 5.2.1.

1. Tomemos $r \geq 1$, fijo arbitrario.

Sea $t_{r-1} := \min\{t \in S : t \geq r-1\}$ y $t_r := \min\{t \in S : t \geq r\}$, es claro que $t_{r-1} \leq t_r$ y por definición $\psi_{r-1} = \psi_{r-1}^{t_{r-1}}$ y $\psi_r = \psi_r^{t_r}$.

Tenemos dos casos, si $t_{r-1} \geq r$, entonces por su definición como mínimo, se tiene: $t_{r-1} = t_r$, luego $\psi_{r-1} = \psi_{r-1}^{t_{r-1}} = \psi_{r-1}^{t_r}$. Si en cambio $t_{r-1} = r-1$ entonces por la propiedad del conjunto S para todo $T \geq r-1$: $\psi_{r-1}^T = \psi_{r-1}^{r-1}$, en particular $\psi_{r-1}^{t_r} = \psi_{r-1}^{r-1} = \psi_{r-1}$. Así en cualquiera de los dos casos se tiene, $\psi_{r-1} = \psi_{r-1}^{t_r}$.

Usando la igualdad $(2)_{t_r}$,

$$\psi_{r-1}^{t_r} = \nabla_{x_r} f_r^0(\hat{x}_r, \hat{u}_r) + \psi_r^{t_r} \circ D_{x_r} f_r(\hat{x}_r, \hat{u}_r) = \nabla_{x_r} f_r^0(\hat{x}_r, \hat{u}_r) + \psi_r \circ D_{x_r} f_r(\hat{x}_r, \hat{u}_r),$$

de donde para todo $r \geq 1$,

$$\psi_{r-1} = D_{x_r} H_r(\hat{x}_r, \hat{u}_{r+1}, \psi_{r+1}).$$

2. Finalmente probemos la segunda implicación. Para $t \geq 0$, sea $s_t := \min\{s \in S : s \geq t\}$, entonces por definición $\psi_t = \psi_t^{s_t}$. Con esto, en la igualdad $(2)_{s_t}$ tenemos

$$0 = \nabla_{u_t} f_t^0(\hat{x}_t, \hat{u}_t) + \psi_t^{s_t} \circ D_{u_t} f_t(\hat{x}_t, \hat{u}_t) = D_{u_t} H_t(\hat{x}_t, \hat{u}_t, \psi_t),$$

con lo cual queda demostrada la implicación **2.** ■

5.3 Otros Principios Débiles Infinitos

Al igual que el Principio de Pontryaguin Lagrange Infinito (Teorema 5.2.1), en esta sección daremos ciertas condiciones al problema $\mathcal{P}(\hat{x}_0)$ de modo que podamos generalizar dicho resultado. Los resultados de esta sección son estudiados en [6].

En esta dirección un resultado que nos ayudará mucho es el siguiente lema, para el cual usaremos una notación estandar, cuando $m, n \in \mathbb{N}$, con $m \leq n$, pondremos $[m,n]_N := [m,n] \cap \mathbb{N}$, y $[m,\infty)_N := [m,\infty) \cap \mathbb{N}$.

Lema 5.3.1 (Lema de Compacidad) *Sea Z un espacio vectorial normado sobre $\mathbb{R}$ de dimensión finita. Para cada $(t,T) \in \mathbb{N}_* \times \mathbb{N}_*$, $\quad t \leq T$ consideremos un elemento $z_t^T \in X$.*

Asumamos que para todo $t \in \mathbb{N}_*$, la secuencia $(z_t^t, z_t^{t+1}, z_t^{t+2}, \cdots)$ es acotada.

Entonces existe una función creciente $\beta : \mathbb{N}_* \to \mathbb{N}_*$, tal que para todo $t \in \mathbb{N}_*$ existe $\bar{z}_t \in X$, que satisface

$$\lim_{T\to\infty} z_t^{\beta(T)} = \bar{z}_t.$$

Prueba. X tiene dimensión finita y por hipótesis $(z_1^T)_{T\geq 1}$ es acotado, entonces del Teorema de Bolzano-Weierstrass existe $\bar{z}_1 \in X$, y una función creciente $\alpha_1 : [1,\infty)_N \to [1,\infty)_N$ tal que

$$\lim_{T\to\infty} z_1^{\alpha_1(T)} = \bar{z}_1.$$

Similarmente para la secuencia $(z_2^{\alpha_1(T)})_{T\geq 2}$ existe $\bar{z}_2 \in X$ y una función creciente $\alpha_2 : [2,\infty)_N \to [2,\infty)_N$ tal que

$$\lim_{T\to\infty} z_2^{\alpha_1(\alpha_2(T))} = \bar{z}_2.$$

Continuando con este proceso obtenemos para cada $t \in \mathbb{N}_*$, una función creciente $\alpha_t : [t,\infty)_N \to [t,\infty)_N$, y un vector $\bar{z}_t \in Z$ tales que

$$\lim_{T\to\infty} z_t^{\alpha_1\circ\alpha_2\circ\ldots\circ\alpha_t(T)} = \bar{z}_t.$$

Definamos la función $\beta : \mathbb{N}_* \to \mathbb{N}_*$, como, $\beta(t) := \alpha_1 \circ \alpha_2 \circ \ldots \circ \alpha_t(t)$, y definamos también las funciones $\delta_t : [t,\infty)_N \to [t,\infty)_N$,

$$\delta_t(T) := \begin{cases} T, & \text{si } T = t, \\ \alpha_{t+1} \circ \ldots \circ \alpha_T(T), & \text{si } T \geq t+1. \end{cases}$$

Con estas definiciones las funciones δ_t son crecientes, en efecto para $T \geq t$

$$\begin{aligned} \delta_t(T+1) &= (\alpha_{t+1} \circ \cdots \circ \alpha_{T-1}) \circ \alpha_T \circ \alpha_{T+1}(T+1) \geq \\ &\geq (\alpha_{t+1} \circ \cdots \circ \alpha_{T-1}) \circ \alpha_T(T+1) > \\ &> (\alpha_{t+1} \circ \cdots \circ \alpha_{T-1}) \circ \alpha_T(T) = \delta_t(T). \end{aligned}$$

Para un valor $t \in \mathbb{N}_*$ fijo podemos expresar $\beta = (\alpha_1 \circ \alpha_2 \circ \cdots \circ \alpha_t) \circ \delta_t$, por ello $(z_t^{\beta(T)})_{T \geq t}$ es una subsecuencia de $(z_t^{\alpha_1 \circ \alpha_2 \circ \cdots \circ \alpha_t(T)})_{T \geq t}$.

Por lo tanto para todo $t \in \mathbb{N}_*$,

$$\lim_{T \to \infty} z_t^{\beta(T)} = \bar{z}_t.$$

■

Finalmente digamos que para poder demostrar los siguientes teoremas hemos de pedir que se satisfaga la siguiente hipótesis.

- Para todo $t \geq 0$, y $(x_t, u_t) \in X_t \times U_t$,

$$D_{x_t} f_t(x_t, u_t) \quad \text{es inversible.} \qquad (Inv)$$

5.3.1 Principio de Pontryagin-Boltyanskii Infinito

Teorema 5.3.1 (Principio Pontryagin-Boltyanskii Infinito (PdPB)$_\infty$) *Para el problema $\mathcal{P}(\hat{x}_0)$ asumamos las hipótesis $(H1)_B$, $(H2)_B$ e (Inv). Sea la sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$, con su camino asociado $(\hat{x}_0, \hat{x}_1, \cdots)$, una* **solución óptima** *que satisface la hipótesis $(I)_B$.*

Entonces existen $a_0 \in \mathbb{R}$ y vectores $\psi_0, \psi_1, \cdots \in \mathbb{R}^n$ tales que

1. $a_0 \geq 0$.

2. (a_0, ψ_0) *es no nulo.*

3. *Para todo $t \geq 1$,*

$$\psi_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t).$$

4. *Para cada $t \geq 0$, se cumple: Para todo $\delta w_t \in L_t$,*

$$\langle \delta w_t \ ; \ \nabla_{u_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t) \rangle \leq 0,$$

donde L_t es la cúpula de $\hat{u}_t$, y $H_t(x_t, u_t, a_0, \psi_t) := a_0 f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$.

Prueba. La sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$ es una solución óptima, entonces del Lema 5.2.1, para cada $T \geq 0$ la sucesión $(\hat{u}_0, \cdots, \hat{u}_T)$, es solución óptima del problema $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$. Para cada uno de estos problemas aplicamos el Corolario 3.5.1.

Por lo tanto para todo $T \geq 1$ tenemos asociados: $a_0^T \in \mathbb{R}$ y vectores $\psi_0^T, \cdots, \psi_{T-1}^T \in \mathbb{R}^n$ tales que:

$(1)_T$ $\quad a_0^T \geq 0$.

$(2)_T$ $(a_0^T, \psi_0^T, \cdots, \psi_{T-1}^T)$ no nulo.

$(3)_T$ Para todo $t = 1, \cdots, T-1$,

$$\psi_{t-1}^T = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t^T). \tag{$(3)_T$}$$

$(4)_T$ Para todo $\delta w_t \in L_t$:

$$\langle \delta w_t \, ; \, \nabla_{u_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t^T) \rangle \leq 0, \tag{$(4)_T$}$$

donde $t = 0, \cdots, T-1$, y L_t es la cúpula de $\hat{u}_t$.

Por su parte $D_{x_t} f_t(\hat{x}_t, \hat{u}_t)$ es invertible, por ello despejando ψ_t^T en las igualdades $(3)_T$ tenemos para $t = 1, \cdots, T-1$,

$$\psi_t^T = -a_0^T \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1} + \psi_{t-1}^T \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1}. \tag{ϕ}$$

Para $t \geq 0$ definamos las siguientes sucesiones

$$(\psi_t^T)_{T \geq t} := (\psi_t^t, \psi_t^{t+1}, \psi_t^{t+2}, \cdots),$$

a partir de éstas construiremos los vectores ψ_t. En este sentido, lo primero es notar que $(a_0^T, \psi_0^T) \neq (0,0)$, pues de lo contrario desde (ϕ) tendríamos

$$0 = \psi_0^T = \psi_1^T = \psi_2^T = \cdots = \psi_{T-1}^T,$$

lo que contradice $(2)_T$, por ello tenemos que $(a_0^T, \psi_0^T) \neq (0,0)$.

En segundo lugar redefinamos las variables adjuntas para $t = 0, \ldots, T-1$,

$$a_0^T := \frac{a_0^T}{\|(a_0^T, \psi_0^T)\|}, \qquad \text{y} \qquad \psi_t^T := \frac{\psi_t^T}{\|(a_0^T, \psi_0^T)\|}.$$

Para estas variables redefinidas siguen siendo ciertas $(3)_T$, $(4)_T$ y (ϕ). Con estos cambios tenemos $\|(a_0^T, \psi_0^T)\| = 1$, en consecuencia $\|a_0^T\| \leq 1$ y $\|\psi_0^T)\| \leq 1$.

Usando (ϕ) y tomando fijo $t = 0$, para $T \geq 1$,

$$\psi_1^T = -a_0^T \nabla_{x_1} f_1^0(\hat{x}_1, \hat{u}_1) \circ [D_{x_1} f_1(\hat{x}_1, \hat{u}_1)]^{-1} + \psi_0^T \circ [D_{x_1} f_1(\hat{x}_1, \hat{u}_1)]^{-1},$$

por lo cual para $T \geq 1$,

$$\|\psi_1^T\| \leq C_1 := \|\nabla_{x_1} f_1^0(\hat{x}_1, \hat{u}_1) \circ [D_{x_1} f_1(\hat{x}_1, \hat{u}_1)]^{-1}\| + \|\,[D_{x_1} f_t(\hat{x}_1, \hat{u}_1)]^{-1}\|.$$

Así pues, el conjunto $\{\psi_1^1, \psi_1^2, \psi_1^3 \cdots\}$ esta acotado por C_1.

Con un razonamiento similar, definiendo inductivamente la constante C_{t-1} tenemos en general para "t" fijo y arbitrario, que para todo $T \geq t$,

$$\|\psi_t^T\| \leq C_t := \|\nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1}\| + C_{t-1}\|\,[D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1}\|,$$

por lo cual el conjunto $\{\psi_t^t, \psi_t^{t+1}, \psi_t^{t+2}, \cdots\}$ está acotado por C_t.

Luego podemos aplicar el lema de compacidad a estas sucesiones por tanto existe una subsucesión $\beta(T)$ y vectores $\{\psi_0, \psi_1, \psi_3, \cdots\}$, para los cuales se cumple

$$\lim_{T\to\infty} \psi_t^{\beta(T)} = \psi_t.$$

Esta subsucesión se puede suponer de modo que $(a_0^{\beta(T)})_{T\geq 0}$ sea también convergente, y tomemos $a_0 := \lim_{T\to\infty} a_0^{\beta(T)}$.

Probemos finalmente que a_0 y estos vectores $\psi_0, \psi_1, \cdots$ satisfacen las implicaciones del teorema.

1. $a_0^{\beta(T)} \geq 0$ entonces $a_0 \geq 0$.

2. Por la construcción realizada tenemos $\|(a_0^{\beta(T)}, \psi_0^{\beta(T)})\| = 1$, luego tomando límites $\|(a_0, \psi_0)\| = 1$.

3. Tomando límites en (ϕ) deducimos para $t \geq 1$

$$\psi_t = -a_0 \nabla_{x_t} f_t^0(\hat{x}_t, \hat{u}_t) \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1} + \psi_{t-1} \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1},$$

por lo cual para $t \geq 1$,

$$\psi_{t-1} = \nabla_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t).$$

4. Finalmente tomando límites, de las igualdades $(4)_{\beta(T)}$ tenemos la última implicación. ■

5.3.2 Principio de Pontryagin-Clarke Infinito

Teorema 5.3.2 (Principio de Pontryagin-Clarke Infinito (PdPC)$_\infty$) *Para el problema $\mathcal{P}(\hat{x}_0)$ asumamos las hipótesis $(H1)_C - (H3)_C$ e (Inv). Sea la sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$, con su proceso asociado $(\hat{x}_0, \hat{x}_1, \cdots)$, una* **solución óptima** *que satisface la hipótesis $(I)_C$.*

Entonces existen $a_0 \in \mathbb{R}$, y vectores $\psi_1, \psi_2, \cdots \in \mathbb{R}^n$ tales que:

1. $a_0 \geq 0$.

2. (a_0, ψ_0) *no nulo.*

3. *Para todo $t \geq 1$,*

$$\psi_{t-1} \in \partial_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t).$$

4. *Para todo $t \geq 0$,*

$$\partial_{u_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t) \cap N_{U_t}(\hat{u}_t) \neq \emptyset,$$

donde $H_t(x_t, u_t, a_0, \psi_t) := a_0 f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$.

Prueba. Desde el Lema 5.2.1, para cada $T \geq 0$ la sucesión $(\hat{u}_0, \cdots, \hat{u}_T)$, es solución óptima del problema $\mathcal{P}_T(\hat{x}_0, \hat{x}_T)$. Para cada uno de estos problemas aplicamos el Corolario 3.6.1.

Por ello existen $a_0^T \in \mathbb{R}$ y vectores $\psi_0^T, \cdots, \psi_{T-1}^T \in \mathbb{R}^n$ tales que:

$(1)_T$ $a_0^T \geq 0$

$(2)_T$ $(a_0^T, \psi_1^T, \cdots, \psi_{T-1}^T)$ no nulo.

$(3)_T$ Para cada $t = 1, \cdots, T-1$, existe $\varphi_t^T \in \partial_{x_t} f_t^0(\hat{x}_t, \hat{u}_t)$, tal que

$$\psi_{t-1}^T = a_0^T \varphi_t^T + \psi_t^T \circ D_{x_t} f_t(\hat{x}_t, \hat{u}_t).$$

$(4)_T$ Para cada $t = 1, \cdots, T-1$, existe $\phi_t^T \in \partial_{u_t} f_t^0(\hat{x}_t, \hat{u}_t)$, tal que para todo $\delta w_t \in T_{U_t}(\hat{u}_t)$,

$$\langle a_0 \phi_t^T + \psi_t \circ D_{u_t} f_t(\hat{x}_t, \hat{u}_t) \ ; \ \delta w_t \rangle \leq 0,$$

donde $N_{U_t}(\hat{u}_t)$ es el cono normal de U_t en $\hat{u}_t$.

Por su parte $D_{x_t} f_t(\hat{x}_t, \hat{u}_t)$ es invertible, por ello despejando ψ_t^T en las igualdades $(3)_T$ tenemos para $t = 1, \cdots, T-1$,

$$\psi_t^T = -a_0^T \varphi_t^T \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1} + \psi_{t-1}^T \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1}. \qquad (\alpha)$$

Para $t \geq 0$ definamos las siguientes sucesiones

$$(\psi_t^T)_{T \geq t} := (\psi_t^t, \psi_t^{t+1}, \psi_t^{t+2}, \cdots),$$

a partir de éstas construiremos los vectores ψ_t. En este sentido, lo primero es notar que $(a_0^T, \psi_0^T) \neq (0,0)$, pues de lo contrario desde (α) tendríamos

$$0 = \psi_0^T = \psi_1^T = \psi_2^T = \cdots = \psi_{T-1}^T,$$

lo que contradice $(2)_T$, por ello tenemos que $(a_0^T, \psi_0^T) \neq (0,0)$.

En segundo lugar redefinamos las variables adjuntas para $t = 0, \ldots, T-1$,

$$a_0^T := \frac{a_0^T}{\|(a_0^T, \psi_0^T)\|}, \qquad \text{y} \qquad \psi_t^T := \frac{\psi_t^T}{\|(a_0^T, \psi_0^T)\|}.$$

Para estas variables redefinidas siguen siendo ciertas $(3)_T$, $(4)_T$ y (α). Con estos cambios tenemos $\|(a_0^T, \psi_0^T)\| = 1$, en consecuencia $\|a_0^T\| \leq 1$ y $\|\psi_0^T)\| \leq 1$.

Usando (α) y tomando fijo $t = 0$, para $T \geq 1$,

$$\psi_1^T = -a_0^T \varphi_1^T \circ [D_{x_1} f_1(\hat{x}_1, \hat{u}_1)]^{-1} + \psi_0^T \circ [D_{x_1} f_1(\hat{x}_1, \hat{u}_1)]^{-1},$$

por lo cual para $T \geq 1$,

$$\|\psi_1^T\| \leq C_1 := c_1 \| \, [D_{x_1} f_1(\hat{x}_1, \hat{u}_1)]^{-1} \, \| + \| \, [D_{x_1} f_t(\hat{x}_1, \hat{u}_1)]^{-1} \, \|,$$

donde c_1 es una constante que acota el conjunto compacto $\varphi_1^T \in \partial_{x_1} f_1^0(\hat{x}_1, \hat{u}_1)$. Así pues, el conjunto $\{\psi_1^1, \psi_1^2, \psi_1^3 \cdots\}$ esta acotado por C_1.

Con un razonamiento similar, definiendo inductivamente la constante C_{t-1} tenemos en general para "t" fijo y arbitrario, que para todo $T \geq t$,

$$\|\psi_t^T\| \leq C_t := c_t \| \, [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1} \, \| + C_{t-1} \| \, [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1} \, \|,$$

donde c_t es la constante que acota al conjunto compacto $\varphi_t^T \in \partial_{x_t} f_t^0(\hat{x}_t, \hat{u}_t)$, por lo cual el conjunto $\{\psi_t^t, \psi_t^{t+1}, \psi_t^{t+2}, \cdots\}$ está acotado por C_t.

Luego podemos aplicar el lema de compacidad a estas sucesiones por tanto existe una subsucesión $\beta(T)$ y vectores $\{\psi_0, \psi_1, \psi_3, \cdots\}$, para los cuales se cumple

$$\lim_{T \to \infty} \psi_t^{\beta(T)} = \psi_t.$$

Esta subsucesión se puede suponer de modo que $(a_0^{\beta(T)})_{T \geq 0}$ sea también convergente, y tomemos $a_0 := \lim_{T \to \infty} a_0^{\beta(T)}$.

Observemos finalmente que para $t \geq 1$,

$$\{\varphi_t^{\beta(1)}, \varphi_t^{\beta(2)}, \cdots\} \subset \{\varphi_t^t, \varphi_t^{t+1}, \ldots\} \subset \partial_{x_t} f_t(\hat{x}_t, \hat{u}_t) \quad \text{(Compacto)},$$
$$\{\phi_t^{\beta(1)}, \phi_t^{\beta(2)}, \cdots\} \subset \{\phi_t^t, \phi_t^{t+1}, \cdots\} \subset \partial_{u_t} f_t(\hat{x}_t, \hat{u}_t) \quad \text{(Compacto)},$$

por ello, desde el lema de compacidad (Lema 5.3.1), es posible tomar una subsucesión de β, digamos γ, tal que para cada $t \geq 1$ existen los límites

$$\varphi_t := \lim_{T \to \infty} \varphi^{\gamma(T)}, \qquad \text{y} \qquad \phi_t := \lim_{T \to \infty} \phi^{\gamma(T)}.$$

Para concluir demostremos que los vectores que hemos construído satisfacen las implicaciones de nuestro teorema.

1. Puesto que $a_0^{\gamma(T)} \geq 0$, tenemos $a_0 \geq 0$.

2. Por la construcción realizada tenemos $\|(a_0^{\gamma(T)}, \psi_0^{\gamma(T)})\| = 1$, luego en el límite $\|(a_0, \psi_0)\| = 1$.

3. Tomando límites en (α) deducimos para $t \geq 1$,

$$\psi_t = -a_0 \, \varphi_t \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1} + \psi_{t-1} \circ [D_{x_t} f_t(\hat{x}_t, \hat{u}_t)]^{-1},$$

por lo cual para $t \geq 1$,

$$\psi_{t-1} \in \partial_{x_t} H_t(\hat{x}_t, \hat{u}_t, a_0, \psi_t).$$

4. Finalmente tomando límites, de las igualdades $(4)_{\gamma(T)}$, obtenemos la última implicación. ■

5.4 Principio Fuerte de Pontryagin Afín Infinito

Concluyamos este capítulo generalizando el Principio Fuerte del Máximo de Pontryagin Afín (Teorema 4.1.1).

Para lograr demostrar este teorema debemos exigir en el problema de horizonte infinito $\mathcal{P}(\hat{x}_0)$ las siguientes hipótesis.

Hipótesis de Inclusión:

- Para todo $t \geq 0$,
$$f_t(X_t \times U_t) \subseteq X_{t+1}. \qquad (Inclu)_\infty$$

Hipótesis de Convergencia:

Para cada $t \geq 0$, y $x_t \in X_t$, (fijo arbitrario) se cumple:

Para todo $(u_t, u_{t+1}, u_{t+2}, \cdots) \in \prod_{s \geq t} U_s$

$$B^t(x_t, u_t, u_{t+1}, u_{t+2}, \cdots) \quad \text{converge en } \mathbb{R}. \qquad (Conver)_\infty$$

Funciones Asociadas a una Sucesión de Decisiones Dada una sucesión de decisiones $(\hat{u}_0, \hat{u}_1, \cdots)$ con su proceso asociado $(\hat{x}_0, \hat{x}_1, \cdots)$, definimos las funciones $g_t : X_t \to \mathbb{R}$,

$$g_t(x_t) := B^t(x_t, \hat{u}_t, \hat{u}_{t+1}, \cdots).$$

A continuación demostremos un lema que no será útil para nuestro teorema.

Lema 5.4.1 (Convergencia de Funciones Afínes) *Para $t \geq 0$ consideremos la sucesión de funciones afines $g_t : X \in \mathbb{R}^n \to \mathbb{R}$, tales que satisfacen las siguientes hipótesis:*

1. *Existen "$n+1$" elementos $x_0, \cdots, x_n \in X$ tales que el conjunto de "n" vectores $\{y_k := x_k - x_{k-1} : k = 1, \cdots, n\}$ es* **linealmente independiente**.
2. *Para cada $k = 0, \cdots, n$, se tiene*
$$\lim_{t \to \infty} g_t(x_k) = L_k \in \mathbb{R}.$$

Entonces la sucesión de funciones $(g_t)_{t \geq 0}$ converge puntualmente a una función afín $g : X \to \mathbb{R}$.

Prueba. Por definición de una función afín, sabemos que para cada $t \geq 0$, podemos expresar las funciones g_t como sigue

$$g_t(x) = \langle \hat{a}_t, x \rangle + \hat{b}_t,$$

donde $\hat{a}_t \in \mathbb{R}^n$ y $\hat{b}_t \in \mathbb{R}$.

Para $t \geq 0$, fijo y arbitrario, consideremos el siguiente sistema de ecuaciones

$$\begin{array}{rcl} g_t(x_1) - g_t(x_0) & = & \langle x_1 - x_0, a\rangle, \\ g_t(x_2) - g_t(x_0) & = & \langle x_2 - x_0, a\rangle, \\ \cdots & \cdots & \cdots \\ g_t(x_n) - g_t(x_0) & = & \langle x_n - x_0, a\rangle, \end{array} \qquad (\phi)$$

donde $a \in \mathbb{R}^n$ es la incógnita. Matricialmente puede expresarse según

$$G_x^t = M_x \cdot a,$$

donde $G_x^t := (g_t(x_1) - g_t(x_0), \cdots, g_t(x_n) - g_t(x_0))^T \in \mathbb{R}^{n\times 1}$ y $M_x \in \mathbb{R}^{n\times n}$ es la matriz cuya fila k-ésima es el vector $x_k - x_0$.

Claramente una solución de este sistema es el vector $\hat{a}_t$.

Debido a que los vectores $x_k - x_0$ son **l.i**, el rango según filas de la matriz M_x es "n", por lo que es una matriz invertible. En consecuencia el sistema (ϕ) tiene por única solución solución al vector $\hat{a}_t$. Por lo que para $t \geq 0$,

$$(M_x)^{-1} \cdot G_x^t = \hat{a}_t. \qquad (\beta)$$

De otro lado del ítem 2. la sucesión $(G_x^t)_{t\geq 0}$ es convergente, por lo cual desde (β) también converge la sucesión $(\hat{a}_t)_{t\geq 0}$, obtenemos:

$$\hat{a} = \lim_{t\to\infty} \hat{a}_t = (M_x)^{-1} \cdot (\lim_{t\to\infty} G_x^t) = (M_x)^{-1} \cdot L_x,$$

con $L_x := (L_1 - L_0, \cdots, L_n - L_0)$.

Por su parte tomando $x \in X$ fijo arbitrario, y expresando $\hat{b}_t = g_t(x) - \langle \hat{a}_t, x\rangle$, vemos que también converge la sucesión $(\hat{b}_t)_{t\geq 0}$, pongamos $\hat{b} := \lim_{t\to\infty} \hat{b}_t$.

Definiendo $g : X \to \mathbb{R}$,

$$g(x) = \langle \hat{a}, x\rangle + b.$$

No es difícil probar que la sucesión de funciones $(g_t)_{t\geq 0}$ converge puntualmente a una función afín $g : X \to \mathbb{R}$. ∎

Observación. No se puede debilitar la hipótesis de este Lema, reduciendo el número de puntos en los que converge.

En vista del resultado anterior es preciso añadir una **hipótesis** a la que denotaremos por $(Form)_X$, más a nuestro problema $\mathcal{P}(\hat{x}_0)$ la cual será **sobre la forma de los conjuntos** X_t.

> Para cada $t \geq 0$ el conjunto X_t existen "$n+1$" elementos $x_0, \cdots, x_n \in X_t$ tales que el conjunto de "n" vectores $\{y_k := x_k - x_{k-1} : k = 1, \cdots, n\}$ es **linealmente independiente**.

Lema 5.4.2 *Para el problema* $\mathcal{P}(\hat{x}_0)$ *asumamos las hipótesis* $(Inclu)_\infty$*,* $(Conver)_\infty$*,* $(HA)_\infty$ *y* $(Form)_X$*. Sea* $(\hat{u}_0, \hat{u}_1, \cdots)$*, una sucesión cualesquiera, con su proceso asociado* $(\hat{x}_0, \hat{x}_1, \cdots)$*. Entonces sus funciones asociadas* $(g_t)_{t\geq 0}$ *son afines.*

Prueba. Dado $t \geq 0$, fijo y arbitrario, probemos que g_t es una función afín. En efecto, definamos para $T \geq t$ las funciones $S_t^T : X_t \to \mathbb{R}$,

$$S_t^T(x_t) := B_T^t(x_t, \hat{u}_t, \cdots, \hat{u}_T).$$

Es claro que para todo $x_t \in X_t$, $\lim_{T\to\mathbb{R}} S_t^T(x_t) = g_t(x_t)$.

Por su parte desde la hipótesis $(HA)_\infty$ las funciones S_t^T son afines. De la hipótesis $(Conver)_\infty$, para todo punto $x_t \in X_t$ las sucesiones $(S_t^T(x_t))_{T\geq t}$ son convergentes en $\mathbb{R}$.

Finalmente de la hipótesis $(Form)_\infty$ existen "n+1" vectores $x_0, \cdots, x_n \in X$ tales que el conjunto de "n" vectores $\{y_k := x_k - x_{k-1} : k = 1, \cdots, n\}$ es **linealmente independiente**.

Por lo dicho anteriormente, en particular para estos vectores las sucesiones $(S_t^T(x_i))_{T\geq t}$ son convergentes en $\mathbb{R}$. Por lo que se cumplen todas las condiciones del Lema 5.4.1. Consecuentemente deducímos que las funciones S_t^T convergen a funciones afines, por lo que las funciones g_t son afines. ■

Estamos en condiciones de formular el Principio Fuerte de Pontryagin Infinito.

Teorema 5.4.1 (PFP Afín Infinito (PFP-Afin$_\infty$)) *Para el problema* $\mathcal{P}(\hat{x}_0)$ *asumamos las hipótesis* $(Inclu)_\infty$*,* $(Conver)_\infty$*,* $(HA)_\infty$ *y* $(Form)_X$*.*

Sea una sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$ *con su proceso asociado* $(\hat{x}_0, \hat{x}_1, \cdots)$*, una* **solución óptima**.

Entonces existen vectores $\hat{\psi}_0, \hat{\psi}_1, \cdots \in \mathbb{R}^n$ *tales que:*

1. *Para* $t \geq 1$,

$$\hat{\psi}_{t-1} = \nabla_x H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t). \qquad (EA)_\infty$$

2. *Para* $t \leq 0$, $\hat{u}_t$ *es* **solución** *del subproblema*

$$\max_{u\in U_t} H_t(\hat{x}, u, \hat{\psi}_t), \qquad (PM)_t$$

donde $H_t(x_t, u_t, \psi_t) = f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t)\rangle$ *es el hamiltoniano.*

Prueba. La prueba es una extensión natural de la que se realiza en el Teorema 4.1.1. Para la sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$ con su proceso asociado $(\hat{x}_0, \hat{x}_1, \cdots)$, definamos sus funciones asociadas $g_t : X_t \to \mathbb{R}$,

$$g_t(x_t) := B^t(x_t, \hat{u}_t, \hat{u}_{t+1}, \cdots).$$

Del Lema 5.4.2 las funciones g_t son afines, en consecuencia diferenciable en todo su dominio. Además desde su definición se ve que las funciones g_t satisfacen la siguiente identidad para $x_t \in X_t$, con $t \geq 0$,

$$g_t(x_t) = f_t^0(x_t) + g_{t+1}(f_t(x_t, \hat{u}_t)). \qquad (\phi)$$

Definamos para $t \geq 1$ los siguientes vectores

$$\hat{\psi}_{t-1} := \nabla_{x_t} g_t(\hat{x}_t).$$

Entonces desde (ϕ) se deduce para $t \geq 1$,

$$\begin{array}{rcl} \nabla_{x_t} g_t(x_t) & = & \nabla_{x_t}[f_t^0(x_t) + g_{t+1}(f_t(x_t, \hat{u}_t))], \\ \nabla_{x_t} g_t(x_t) & = & \nabla_{x_t} f_t^0(x_t) + \nabla_{x_{t+1}} g_{t+1}(f_t(x_t, \hat{u}_t)) \circ D_{x_t}(f_t(x_t, \hat{u}_t)). \end{array} \qquad (\beta)$$

Usando la definición del hamiltoniano $H_t(x_t, u_t, \psi_t) = f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$, de la identidad (β) obtenemos $(EA)_\infty$.

Finalmente desde la optimalidad de la sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$ para el problema $\mathcal{P}(\hat{x}_0)$, es también óptima la sucesión $(\hat{u}_t, \hat{u}_{t+1}, \cdots)$ para el problema $\mathcal{P}^t(\hat{x}_t)$, por lo que para toda $u_t \in U_t :\ B_T(\hat{x}_t, u_t, \hat{u}_{t+1}, \cdots) \leq B_T(\hat{x}_t, \hat{u}_t, \hat{u}_{t+1}, \cdots)$. Entonces $\hat{u}_t$ es solución del problema

$$\max_{u \in U_t} B_T(\hat{x}_t, u_t, \hat{u}_{t+1}, \cdots) = g_t(\hat{x}_t).$$

A partir de la hipótesis $(HA)_\infty$ deducimos que las funciones g_t son afínes. Por lo tanto para todo $x \in X_t$ se cumple

$$g_t(x) - g_t(\hat{x}_t) = \langle \nabla g_t(\hat{x}_t); x - \hat{x}_t \rangle.$$

Con este resultado y de la definición de g_t tenemos

$$\begin{array}{rcl} g_t(\hat{x}_t) & = & \max_{u_t \in U_t}[f_t^0(\hat{x}_t, u_t) + B_T^{t+1}(f_t(\hat{x}_t, u_t); \hat{u}_{t+1}, \cdots, \hat{u}_{T-1})] \\ & = & \max_{u_t \in U_t}[f_t^0(\hat{x}_t, u_t) + g_{t+1}(f_t(\hat{x}_t, u_t))] \\ & = & \max_{u_t \in U_t}[f_t^0(\hat{x}_t, u_t) + \langle \nabla g_{t+1}(\hat{x}_{t+1}); f_t(\hat{x}_t, u_t) - \hat{x}_{t+1} \rangle + g_{t+1}(\hat{x}_{t+1})] \\ & = & C(\hat{x}_t) + \max_{u_t \in U_t}[f_t^0(\hat{x}_t, u_t) + \langle \nabla g_{t+1}(\hat{x}_{t+1}); f_t(\hat{x}_t, u_t) \rangle], \end{array}$$

donde $C(\hat{x}_{t+1}) = g_{t+1}(\hat{x}_{t+1}) - \langle \nabla g_{t+1}(\hat{x}_{t+1}); \hat{x}_{t+1} \rangle$.

Desde su definición $\hat{\psi}_t = \nabla g_{t+1}(\hat{x}_{t+1})$, y teniendo en cuenta el Hamiltoniano.

$$\begin{array}{rcl} g_t(\hat{x}_t) & = & C(\hat{x}_t) + \max_{u_t \in U_t}[f_t^0(\hat{x}_t, u_t) + \langle \hat{\psi}_t; f_t(\hat{x}_t, u_t) \rangle] \\ & = & C(\hat{x}_t) + \max_{u_t \in U_t} H_t(\hat{x}_t, u_t, \hat{\psi}_t). \end{array}$$

Calculando $g_t(\hat{x}_t) - C(\hat{x}_t)$ tenemos

$$\begin{array}{rcl} g_t(\hat{x}_t) - C(\hat{x}_t) & = & g_t(\hat{x}_t) - g_{t+1}(\hat{x}_{t+1}) + \langle \nabla g_{t+1}(\hat{x}_{t+1}); \hat{x}_{t+1} \rangle) \\ & = & f_t^0(\hat{x}_t, \hat{u}_t) + \langle \hat{\psi}_t, \hat{x}_{t+1} \rangle) \\ & = & H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t). \end{array}$$

Por lo tanto

$$H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t) = \sup_{u \in U_t} H_t(\hat{x}, u, \hat{\psi}_t).$$

Así pues $\hat{u}_t \in U_t$ es una solución óptima del problema $(PM)_t$. ■

El teorema anterior es bastante general, demostremos un corolario más concreto.

Para ello definamos el problema

$$\begin{array}{rl} \max & B(x_0, u_0, u_1, \cdots) := \sum_{t=0}^{\infty} \beta^t f^0(x_t, u_t) \\ \text{s.a} & \text{para todo } \ t \geq 0, \ x_t \in X, \, u_t \in U, \\ & x_{t+1} = f_t(x_t, u_t), \ x_0 = \hat{x}_0 \in X. \end{array} \qquad \mathcal{P}^d(\hat{x}_0)$$

en el que $0 < \beta < 1$ y para todo $t \geq 0$,

- $X \subseteq \mathbb{R}^n$ que satisface las hipótesis $(Form)_X$ y $U \subseteq \mathbb{R}^m$ compacto.
- $f^0 : X \times U \to \mathbb{R}, \ f_t : X \times U \to X$, son funciones afines respecto a la variable de estado "x" y continuas en su respectivo dominio.
- Existe $M \in \mathbb{R}$ tal que para todo $(x, u) \in X \times U : \ |f^0(x, u)| \leq M$.

Hemos denotado a este problema con superíndice "d" haciendo énfasis en el hecho de que se usan descuentos del valor (tasa de descuento β). Para este tipo de problemas tenemos la siguiente proposición que viene a ser un corolario del Teorema 5.4.1 (PFP-Afin$_\infty$).

Proposición 5.4.1 (PFP-Afin$_\infty$ con Tasa de Descuento) *En el problema $\mathcal{P}^d(\hat{x}_0)$ sea la sucesión $(\hat{u}_0, \hat{u}_1, \cdots)$ con su proceso asociado $(\hat{x}_0, \hat{x}_1, \cdots)$ una* **solución óptima.**

Entonces existen vectores $\hat{\psi}_0, \hat{\psi}_1, \cdots \in \mathbb{R}^n$, tales que:

1. *Para $t \geq 1$,*

$$\hat{\psi}_{t-1} = \nabla_x H_t(\hat{x}_t, \hat{u}_t, \hat{\psi}_t). \qquad (EA)_\infty$$

2. *Para $t \leq 0$, $\hat{u}_t$ es* **solución** *del subproblema*

$$\max_{u \in U_t} H_t(\hat{x}, u, \hat{\psi}_t), \qquad (PM)_t$$

donde $H_t(x_t, u_t, \psi_t) = f_t^0(x_t, u_t) + \langle \psi_t, f_t(x_t, u_t) \rangle$ es el hamiltoniano.

Prueba. Desde el Teorema 5.4.1 sólo es necesario probar que el problema $\mathcal{P}^d(\hat{x}_0)$ satisface las hipótesis $(Inclu)_\infty$, $(Conver)_\infty$, $(HA)_\infty$ y $(Form)_X$. Salvo $(Conver)_\infty$, todas éstas se demuestran inmediatamente.

Teniendo en cuenta que la función f^0 es continua y que $f(X \times U) \subseteq X$, se deduce para $S_1 \leq S_2 \in \mathbb{N}$,

$$\left| \sum_{t \geq S_1}^{S_2} \beta^t f^0(x_t, u_t) \right| \leq \sum_{t \geq S_1}^{S_2} \beta^t \left| f^0(x_t, u_t) \right| \leq \left(\sum_{t \geq S_1}^{S_2} \beta^t \right) \cdot M, \qquad (\phi)$$

donde $x_{S_1} \in X_{S_1}$ es fijo arbitrario, y también lo es $(u_t)_{t \geq S_1} \in \prod_{t \geq S_1} U_t$.

No es difícil ver que desde la convergencia de la serie $\sum \beta^t$ y desde (ϕ) se demuestra $(Conver)_\infty$. ■

Capítulo 6

Conclusiones

En esta tesis hemos estudiado las 2 principales aproximaciones que permiten resolver problemas Control Óptimo Discreto.

La Programación Dinámica nos brinda un método estándar y ámpliamente utilizado (Algoritmo de Bellman), para resolver gran variedad de estos problemas. Sin embargo este método exige una gran cantidad de cálculos (requiere resolver todos los subproblemas posibles, antes de arribar a la solución). Es por ello que se requiere buscar métodos alternativos o complementarios.

La aproximación a través de los Principios de Débiles Pontryagin, es un método alternativo, (mucho más eficaz en algunos casos). Hemos logrado realizar un estudio amplio sobre estos principios, exponiéndolos dentro de un marco conceptual que clarifica sus desventajas y su generalidad.

Más aún, observamos desde un punto de vista unificado que todos estos principios provienen de "debilitar" la "discretización del Principio del Máximo de Pontryagin". Esto último nos instó, a investigar en que casos no es necesario "debilitarlo", problema para el cual descubrimos un caso (afinidad en la variable de estado) en el que ésto es posible.

A parte del interés teórico de resolver esta pregunta, la motivación también es de índole práctica. En efecto, en los Principios Débiles las decisiones óptimas $\hat{u}_t$ son puntos críticos de los subproblemas $(PM)_t$. El número de puntos críticos puede ser mucho mayor que el número de soluciones, por lo que asegurar que de hecho $\hat{u}_t$ es solución (Principio Fuerte) reduce el número de candidatos para resolver nuestro problema de Control Óptimo Discreto.

Naturalmente surge la motivación de extender estos resultados para problemas de Horizonte Infinito, lo cual se ha conseguido exigiendo ciertas hipótesis para el problema. Particularmente resaltemos el Principio Fuerte de Pontryagin Afin Infinito, por lograr generalizar de un modo exacto el PFP en este nuevo contexto.

De otro lado, aún cuando esta Tesis esta más abocada a problemas de orden teórico, no podemos dejar de señalar que la Teoría de Control Óptimo de Procesos Discretos tiene abundantes aplicaciones en la Teoría Económica, Macroeconomía Dinámica, Teoría de Juegos, Administración Científica, Investigación de Operaciones, etc. (ver [3, 5, 4, 11, 13, 14, 18, 21, 23])

Quisiera concluir esta Tesis proponiendo ciertas ideas en las que se podría trabajar en lo futuro, extendiendo los resultados del presente trabajo.

En primer lugar realizar una aproximación teórica similar para el Problema de Control Óptimo Discreto Estocástico.

Se nos presenta también la tarea de aplicar a problemas concretos los resultados que aquí hemos obtenido.

Finalmente quisiera proponer la siguiente investigación. Durante la realización de la Tesis hemos visto que es posible dar un marco teórico único para los diferentes Principios Débiles (discretización del Principio del Máximo de Pontryagin), a su vez hemos visto que es posible investigar condiciones que deben satisfacer los problemas para que sus soluciones satisfagan tal o cual condición.

De estas observaciones vemos que en realidad estamos trabajando a un nivel más alto (meta-nivel), es decir, no estamos trabajando con problemas concretos, sino con familias de problemas (las cuales son descritas por alguna característica) Cada familia satisface ciertas condiciones, por lo tanto sus soluciones óptimas cumplen cierto "condición necesaria", en consecuencia tienen cierta "forma de comportarse".

Lo anterior nos hace recordar a la Teoría Cualitativa de las Ecuaciones Diferenciales, en la que no importan las soluciones en sí, sino la forma en que se comportan.

Tenemos pues un problema abierto: Clasificar los diversos problemas de Control Óptimo de Procesos Discretos, y describir el comportamiento de sus soluciones según cada caso. Para realizar por completo esta tarea falta recorrer un largo camino. Los resultados de esta Tesis representan desde esta perspectiva un pequeño paso, de la que podría llamarse "Teoría Cualitativa de Problemas de Control Óptimo de Procesos Discretos."

Apéndice A

Cálculo Generalizado de Clarke

En esta apéndice presentaremos algunos conceptos y resultados desarrollados por Clarke (ver [10, 17]) con los cuales lográ extender la noción de diferenciabilidad y obtener un nuevo cálculo al que se denomina Generalizado.

También se logra extender el Teorema de la Multiplicación de Lagrange pero esta vez para espacios de Banach y funciones Local Lipschitz.

Durante esta sección trabajeremos en el espacio de Banach X, y denotaremos la norma de este como: $\|x\|$

A.1 Gradiente Generalizada de Clarke

Comenzemos recordando las definiciones de las funciones Lipschitz y local lipschitzs

Definición A.1.1 (Funciones Lipschitz) *Sea $Y \subseteq X$ y $f : Y \to \mathbb{R}$ Diremos que f es Lipschitz (de rango K) en $x \in Y$ Si existe $K \geq 0$ tal que: Para todo $y, y' \in Y$ se tiene*

$$|f(y) - f(y')| \leq K\|y - y'\|.$$

Diremos que f es **Lipschitz en** *Y si f es Lipschitz para todo $x \in Y$.*

De otra parte diremos que f es **Local Lipschitz** *en $x \in Y$, si existe $B_\epsilon(x)$ tal que $f|_{B_\epsilon(x)}$ es Lipschitz en $B_\epsilon(x)$*

Finalmente diremos que f es **Local Lipschitz** *en Y, si lo es para $x \in Y$ cualesquiera.*

Notemos que al trabajar con funciones Local Lipschitz estamos en contexto más amplio que el diferenciable y el convexo. En efecto, son resultados conocidos que si f es convexo entonces es Local Lipschitz, y tambien que si f es C^1 entonces es Local Lipschitz.

También definamos un concepto que generaliza el de derivada direccional:

Definición A.1.2 (Derivada Direccional Generalizada) *Sea f Local Lipschitz en x, y $v \in X$, definimos la Derivada Direccional Generalizada de f en x en la dirección v, denotada por $f^\circ(x;v)$, como sigue*

$$f^\circ(x;v) := \limsup_{\substack{x\to y \\ t\downarrow 0}} \frac{f(y+tv)-f(y)}{t}.$$

Equivalente a la definición anterior tenemos, para $\delta > 0$ fijo:

$$f^\circ(x;v) := \lim_{\epsilon\downarrow 0} \sup_{\|x'-x\|<\epsilon\delta} \sup_{0<t<\epsilon} \frac{f(x'+tv)-f(x')}{t}.$$

Proposición A.1.1 *Sea f una función Local Lipschitz en x, de rango K. Entonces se cumple:*

1. *La función $v \mapsto f^\circ(x;v)$ esta bien definida sobre todo X es finita, positivamente homogénea y subaditiva. Satisface además*

$$|f^\circ(x;v)| \leq K\|v\|.$$

2. *$f^\circ(x;v)$ es semicontinua superior como función sobre (x,v) y para x fijo, es Lipschitz de rango K sobre X, es decir: Para todo $v,w \in X$*

$$|f^\circ(x;v) - f^\circ(x;w)| \leq K\|v-w\|$$

3. *$f^\circ(x;-v) = (-f)^\circ(x;v)$*

Definamos ahora el concepto clave de esta teoría.

Definición A.1.3 (Gradiente Generalizada) *Sea f una función Local Lipschitz en x, definiremos la gradiente generalizada de f en x, denotada por $\partial f(x)$*

$$\partial f(x) := \{\zeta \in X^* : f^\circ(x;v) \geq \langle \zeta, v\rangle, \qquad \textit{para toda} \quad v \in X\},$$

donde X^*, es el espacio de funcionales lineales continuas sobre X (espacio dual).

Para uso futuro denotemos $\|\zeta\|_*$, la norma de $\zeta \in X^*$ en el espacio dual X^*

Proposición A.1.2 *Sea f Local Lipschitz de rango K en x. Entonces*

1. $\partial f(x)$ es no vacío, convexo, y $\|\zeta\|_* \leq K$ para todo $\zeta \in \partial f(x)$.

2. Para todo $v \in X$ tenemos

$$f^\circ(x; v) = max\{\langle \zeta, v\rangle : \zeta \in \partial f(x)\}.$$

Resultará útil comparar esta nueva definición con las clásicas, consideremos dos espacios de banach, X e Y, y una función $F : X \to Y$.

Definición A.1.4 (Derivadas Direccionales) *Para $x \in X$ y $v \in X$ definimos la* derivada direccional *de F en x en la dirección v, como*

$$F'(x; v) := \lim_{t\downarrow 0} \frac{F(x + tv) - F(x)}{t}.$$

Diremos que F tiene Derivada Direccional de Hadamard *en x y en la dirección v, denotándola similarmente, si existe el siguiente límite:*

$$F'(x; v) := \lim_{t\downarrow 0} \lim_{h'\to h} \frac{F(x + th') - F(x)}{t}.$$

Definición A.1.5 (Derivadas) *Se dice que F tiene Derivada de Gateaux en el punto $x \in X$. Si existe $T \in L(X, Y)$ tal que para todo v, existe $F'(x; v)$, y además*

$$F'(x; v) = T(v).$$

En este caso se denotará T, como $DF(x)$.

Por su parte se dirá que F tiene Derivada de Hadamard *en $x \in X$, si la función $F'(x, \cdot) : X \to Y$, definida usando la derivada direccional de Hadamard, existe y es lineal.*

Diremos que F es Frechet Diferenciable *en $x \in X$ si es Gateaux diferenciable con derivada $DF(x)$, y cumple:*

- *$DF(x)$ es una funcional continua.*
- *Para toda $h \in X$, $\qquad F(x + h) = F(x) + F'(x, h) + 0(\|h\|)$.*

Finalmente diremos que F es **estrictamente diferenciable** *en $x \in X$, si existe $T \in L(X, Y)$ tal que para toda $v \in X$*

$$\lim_{\substack{x'\to x \\ t\downarrow 0}} \frac{F(x' + tv) - F(x')}{t} = T(v).$$

Denotaremos a esta derivada por $D_sF(x)$

Ahora citemos algunos resultados que muestran las relaciones que existen entre la derivada de Clarke y las otras derivadas.

Proposición A.1.3 *Sea f Local Lipschitz en $x \in X$ y admite la derivada de Gateaux entonces $Df(x) \in \partial f(x)$.*

El mismo resultado se tiene para para la derivada de Hadamard, Frechet o estricta.

Proposición A.1.4 *Si f es estrictamente diferenciable en $x \in X$, entonces f es Local Lipschitz en x y $\partial f(x) = \{D_s f(x)\}$. Reciprocamente, si f es local Lipschitz en x y $\partial f(x)$ tiene un único elemento ζ, entonces f es estrictamente diferenciable en x y $D_s f(x) = \zeta$.*

También tenemos una relación con el subdiferencial del Análisis Convexo.

Proposición A.1.5 *Cuando f es convexo sobre U y local lipschitz en x, entonces $\partial f(x)$ coincide con la subdiferencial en x del análisis convexo, y $f^\circ(x; v)$ coincide con la derivada direccional $f'(x; v)$ para cada v*

A.2 Cálculo Generalizado

Ahora abordaremos algunos aspectos del "Cálculo" que se origina con esta definición de derivada.

Proposición A.2.1 *Para f Local Lipschitz en x, se cumple:*

- *Sea s un escalar fijo y arbitrario, entonces:* $\qquad \partial(sf)(x) = s\,\partial f(x)$.
- *Para un conjunto de escalares $s_1, \cdots, s_n$:* $\partial(\sum_{i=1}^n s_i f_i)(x) \subset \sum_{i=1}^n s_i \partial f_i(x)$.

Este nuevo cálculo se hace más preciso cuando las funciones cumplen una condición adicional, por ello definamos el concepto de funciones regulares.

Definición A.2.1 (Función Regular) *Diremos que f es regular en $x \in X$ si se satisfaze:*

1. *Para todo vector $v \in X$ existe la derivada direccional $f'(x; v)$.*
2. *Para todo vector $v \in X$ se cumple $f'(x; v) = f^\circ(x; v)$.*

En el ítem 2 se tendrá la igualdad si al menos una función es estrictamente diferenciable en x.

Las funciones regulares permiten generalizar algunas propiedades del cálculo.

Proposición A.2.2 *Si cada f_i es regular en x, entonces*

$$\partial(\sum_{i=1}^{n} f_i)(x) = \sum_{i=1}^{n} \partial f_i(x).$$

También se cumple cuando $s_i \geq 0$

$$\partial(\sum_{i=1}^{n} s_i f_i)(x) \subset \sum_{i=1}^{n} s_i \partial f_i(x).$$

Veamos ahora que tan general es el concepto de funciones regulares.

Proposición A.2.3 *Sea f Local Lipschitz en $x \in X$, tenemos:*

a) Si f es estrictamente diferenciable en x, entonces f es regular en x.

b) Si f es convexa, entonces f es regular en x.

c) Sea f_i funciones regulares y $s_i \geq 0$, entonces $\sum_{i=1}^{n} s_i f_i(x)$ es regular.

d) Si f tiene derivada de Gateaux $Df(x)$ y es regular en x, entonces

$$\partial f(x) = \{Df(x)\}.$$

Citemos algunos resultados que generalizan sus análogos para el cálculo usual.

Teorema A.2.1 (Valor Medio) *Sea x e y puntos en X y supongamos que f es Lipschitz en un conjunto abierto conteniendo el segmento $[x, y]$. Entonces extiste un punto $u \in (x, y)$ tal que*

$$f(y) - f(x) \in \langle \partial f(u), y - x \rangle.$$

Proposición A.2.4 (productos) *Sean f_1, f_2 Local Lipschitz en x. Entonces el producto $f_1 f_2$ es Local Lipschitz en x, y se cumple:*

$$\partial(f_1 f_2)(x) \subset f_2(x)\partial f_1(x) + f_1(x)\partial f_2(x).$$

Si adicionalmente tuviesemos $f_1 \geq 0$, $f_2 \geq 0$ y si f_1, f_2 son ambos regulares en x, entonces se cumple la igualdad y el producto $f_1 f_2$ es regular en x.

De otro lado también podemos generalizar las derivadas parciales.

Definición A.2.2 (Gradiente Generalizada Parcial) *Sea $X = X_1 \times X_2$, con X_1, X_2 espacios de Banach, sea $f(x_1, x_2)$ definido sobre X y Local Lipschitz en (x_1, x_2).*

Denotemos a $\partial_1 f(x_1, x_2)$ al gradiente generalizado de $f(\cdot, x_2)$ en x_1, y por $\partial_2 f(x_1, x_2)$ al gradiente generalizado de $f(x_1, \cdot)$ en x_1.

Proposición A.2.5 *Si f es regular en $x = (x_1, x_2)$, entonces*

$$\partial f(x_1, x_2) \subset \partial_1 f(x_1, x_2) \times \partial_2 f(x_1, x_2)$$

Como hemos visto esta nueva aproximación extiende muchos resultados clásicos, ahora veamos como también permite extender nociones geométricas del Análisis Convexo. Tenemos por ejemplo una nueva definición para el cono tangente de un conjunto C. Para ello recordemos la definición de la **función distancia**. Dado $C \subset X$, un conjunto no vacío, entonces definimos la función distacia, $d_C(\cdot) : X \to \mathbb{R}$, según

$$d_C(x) = \inf\{\|x - c\| : c \in C\}.$$

No es difícil probar que esta función es Lipschitz en todo X. Habiendo hecho estas precisiones podemos definir:

Definición A.2.3 (Cono Tangente) *Sea $x \in C$, diremos que un vector $v \in X$ es* **tangente** *a C en x siempre que $d_C^\circ(x; v) = 0$. Al conjunto de todos los vectores tangentes se le denomina* **Cono Tangente** *de C en x y se le denota $T_C(x)$.*

Como consecuencia de la Proposición A.1.1 se tiene que el cono tangente es cerrado convexo y es cono. Definamos además al cono normal:

Definición A.2.4 (Cono Normal) Definimos el Cono Normal como el polar del cono tangente.

$$N_C(x) = \{\zeta \in X^* : \langle \zeta, v \rangle \leq 0, \quad \text{para todo } v \in T_C(x)\}.$$

Podemos calcular en Cono Normal del siguiente modo.

Proposición A.2.6

$$N_C(x) = \mathrm{cl}\,\{\bigcup_{\lambda \geq 0} \lambda\ \partial \mathrm{d_C(x)}\},$$

donde cl *es la cerradura, para la topología $*$, del Dual.*

Es un resultado conocido que si C es convexo, $N_C(x)$ coincide con el cono normal en el sentido del análisis convexo, por ello podemos ver como estos conceptos generalizan la noción del análisis convexo.

A continuación presentemos un resultado que muestra que la noción de *Cono Tangente* no depende de la norma que se use en el espacio X.

Teorema A.2.2 (Caracterización Intrínseca del Cono Tangente) *El cono tangente en x, $T_C(x)$, se puede puede describir de la siguiente manera*

$$T_C(x) = \{v \in X/\forall(t_j) \downarrow 0; \forall x_j \to x \ con \ x_j \in C; \quad \exists v_j \to v : x_j + tv_j \in C\}.$$

A partir del resultado anterior se puede demostrar.

Corolario A.2.1 *Sea $X = X_1 \times X_2$, donde X_1 y X_2 son espacios de Banach, y sea $x = (x_1, x_2) \in C_1 \times C_2$ donde $C_1 \subset X_1, C_2 \subset X_2$, entonces*

$$T_{C_1 \times C_2}(x) = T_{C_1}(x_1) \times T_{C_2}(x_2),$$

$$N_{C_1 \times C_2}(x) = N_{C_1}(x_1) \times N_{C_2}(x_2).$$

Por último definamos los **Conjuntos Regulares** para ello tenemos que definir el concepto de **Cono Contingente** de C en el punto x, denotado $K_C(x)$

$$K_C(x) := \{v \in X : \forall\epsilon > 0, \exists t \in (0, \epsilon) \text{ y } w \in v + \epsilon B : x + tw \in C\}.$$

Desde la definición de $T_C(x)$, se ve que siempre esta incluído en $K_C(x)$.

Definición A.2.5 *El conjunto C es* **regular** *en $x \in X$, cuando $T_C(x) = K_C(x)$.*

Se puede probar que todo conjunto convexo es regular en todo punto.

El concepto de conjunto regular se introduce entre otras razones para poder cracterizar con más precisión los conos tangentes y normales. En este sentido tenemos el siguiente resultado.

Proposición A.2.7 *Sea el conjunto C*

$$C := \{y \in X : f_1(y) \leq 0, f_2(y) \leq 0, \cdots, f_1(n) \leq 0\}.$$

Tomemos además $x \in X$ tal que $f_i(x) = 0 \quad (\forall i = 1, \ldots, n)$ Bajo el supuesto de que cada f_i sea estrictamente diferenciable en x, y si $D_s f_i, \quad i = 1, 2, \ldots, n$, son positivamente lineales independientes. Tendremos que C es regular en x , y tenemos:

$$N_C(x) = \left\{ \sum_{i=1}^{n} \lambda D_s f_i(x) : \lambda_i \geq 0, \ con \ i = 1, 2, \cdots, n \right\}$$

Mencionemos finalmente resultados que tienen que ver con la optimización.

Proposición A.2.8 (Extremo Local) *Si f tiene máximo o mínimo local en x, entonces $0 \in \partial f(x)$.*

Proposición A.2.9 *Sea f una función Lipschitz de rango K sobre el cojunto S. Sea además $x \in C \subset S$ y supongamos que f alcanza su mínimo sobre C en el punto x. Entonces para todo $\hat{K} \geq K$, la función: $g(y) := f(y) + \hat{K} d_C(y)$, alcanza el mínimo de S en el punto x.*

Corolario A.2.2 *Sea f local lipschitz en x y toma su mínimo sobre C en x. Entonces se cumple: $0 \in \partial f(x) + N_C(x)$.*

Finalmente presentamos un resultado fundamental, nos referimos a la generalización del Teorema de Multiplicadores de Lagrange, obtenido por Clarke. Para hacerlo con precisión defininamos los conceptos previos necesarios. En primer lugar mostremos el problema a tratar.

Planteamiento del Problema Sea X un espacio de Banach, y $f : X \to \mathbb{R}$. Consideremos las siguientes restricciones:

- $g_i(x) \leq 0, \quad i = 1, 2, \ldots, n$. Donde $g_i : X \to \mathbb{R}$
- $h_i(x) = 0, \quad i = 1, 2, \ldots, m$. Donde $h_i : X \to \mathbb{R}$
- $x \in C$, donde $C \subset X$

Usando las funciones anteriores definamos las funciones, $g := [g_1, g_2, \ldots, g_n] : X \to \mathbb{R}^n$ y $h := [h_1, h_2, \ldots, h_m] : X \to \mathbb{R}^m$. A su vez, con estas funciones definamos el **Lagrangiano**, $L(x, \lambda, r, s, k) : X \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}$,

$$L(x, \lambda, r, s, k) := \lambda f(x) + \langle r, g(x) \rangle + \langle r, g(x) \rangle + k|(\lambda, r, s)| d_C(x)$$

donde d_C denota la función distancia.

Para el problema anterior supongamos que C sea cerrado y que cada función f, g_i, h_j son Local Lipschitz en C, tenemos el siguiente teorema.

Teorema A.2.3 (Multiplicadores de Lagrange Generalizado) *Sea $x \in X$ un punto optimal de f sujeto a las restricciones anteriores. Entonces para k suficientemente grande, existen $\lambda \geq 0, r \geq 0$ y s* **no todos nulos**, *tales que:*

$$\langle r, g(x) \rangle = 0 \quad y \quad 0 \in \partial_x L(x, \lambda, r, s, k).$$

Apéndice B

Teoría de Optimización de Boltyanskii

B.1 Definiciones Previas

Este Apéndice resume los principales conceptos y teoremas de la Teoría de Optimización de Boltyanskii, [7].

En este apéndice X será un espacio de Banach real de dimensión finita. X^* es su dual.

La Aproximación de Boltyanskii, busca extender las condiciones de óptimo, tratando en primer lugar de generalizar la noción del cono tangente. Es por ello que comienza definiendo, una noción que captura la propiedad más distintivad del cono tangente, es decir la aproximación de primer orden. Observando que el cono tangente es convexo, se definirá entonces de un modo bastante general **la cúpula de un conjunto** en un punto.

Definición B.1.1 (Cúpula de un conjunto) *Dado un conjunto $\Omega \subseteq E^n$ (no necesariamente abierto o convexo) y $q \in \Omega$ diremos que el conjunto C, el cual es cono convexo, con vértice 0 es la* **Cúpula del conjunto** Ω *en el punto q*

Si existe un entorno $B_\epsilon(0)$ y una función continua $f : C \cap B_\epsilon(0) \to X$ tales que:

(1) $f(C \cap B_\epsilon(0)) \subset \Omega$ y $f(0) = q$

(2) Existe $\rho : C \to X$ tal que

$$\lim_{\substack{h \to 0 \\ h \in C}} \rho(h) = 0,$$

y para todo $h \in C \cap B_\epsilon(0)$ tenemos $f(h) = f(0) + h + \|h\| \cdot \rho(h)$.

Recordemos que un cono C de vértice 0, es simplemente un cono, tal como se conoce [i.e un conjunto tal que $\forall a \in C$ se cumple: Si $a \in C$, entonces $t.a \in C \quad \forall t \geq 0$].

Es claro que si $C' \subset C$, es otro cono con vértice 0 entonces tambien es una cúpula del conjunto Ω en el punto q. Por ello esta noción no es unívoca como otras, como el cono tangente o el normal. Pero en cierto sentido lo generaliza. Como vemos en el siguiente resultado.

Proposición B.1.1 *Sea $\Omega \subset X$ convexo, cerrado y $q \in \Omega$ Entonces $T_\Omega(q)$ es cúpula de Ω en el punto q.*

Proposición B.1.2 *Sea $\Omega = \{x \in X : g(x) \leq 0\}$ tal que:*

(i) $g : A \to \mathbb{R}$ *con* $A \subset X$ *, y* $g \in C^1(A; \mathbb{R})$

(ii) $z_0 \in \Omega$ *y* $g(z_0) = 0$ *y* $Dg(z_0) \neq 0$

Entonces: $M := \{h \in X : \langle h, Dg(z_0)\rangle \leq 0\}$ es la cúpula de Ω en el punto z_0.

Para el conjunto $\Omega = \{x \in X : g^1(x) \leq 0, \ldots, g^k(x) \leq 0\}$ y el punto $x_0 \in \Omega$ denotemos como ya es tradicional por $I(x_0)$ al conjunto de restricciones activas, es decir, $I(x_0) := \{j \in [1, \cdots, k] : g^j(x_0) = 0\}$. Definamos ahora una noción que será muy importante en varios teoremas de optimalidad.

Definición B.1.2 (Punto de Boltyanskii no degenerado) *El punto x_0 se dirá que es* **Boltyanskii no degenerado** *siempre que exista $a \in X$ tal que:*

$$\langle a, Dg^i(x_0)\rangle < 0 \qquad \forall i \in I(x_0).$$

Esta definición es esencial para la siguiente proposición

Proposición B.1.3 *Sea $A \subseteq X$ un abierto y para $i \in [1, \cdots, k]$ las funciones $g^i : A \to \mathbb{R}$ tales que $g^i \in C^1(A, \mathbb{R})$. Sea $\Omega := \{z \in X : g^i(z) \leq 0$ con $i \in [1, \cdots, k]\}$. Tomemos un punto $z_0 \in \Omega$ que sea* **no degenerado**.

Entonces el conjunto $C := \{h \in X : \forall i \in I(z_0) : Dg^i(x_0) \cdot h \leq 0\}$ es una cúpula del conjunto Ω en el punto z_0.

Corolario B.1.1 *Sea $\Omega := \{z \in X : g^i(z) = 0$ con $i \in [1, \cdots, k]\}$ tales que $g^i \in C^1(A, \mathbb{R})$. Tomemos un punto $z_0 \in \Omega$ que sea* **no degenerado**.

Entonces

$$C := \{h \in X : \forall i \in I(z_0) : Dg^i(x_0) \cdot h = 0\}$$

es una cúpula del conjunto Ω en el punto z_0.

B.2 Condiciones Necesarias de Optimalidad

Para demostrar las condiciones de optimalidad, en esta aproximación, es necesario demostrar un teorema llamado de intersección.

Definición B.2.1 (Propiedad de separación) *Diremos que una familia de conos, convexos y cerrados* $\{K_1, \ldots, K_s\}$ *de vértice* 0*, cumple la* **Propiedad de Separación.**

Si existen 2 subconjuntos disjuntos I_1, I_2 *tales que* $I_1 \cup I_2 = \{1, \ldots, s\}$ *tales que, para los conjuntos:* $\bar{K} = \bigcap_{i \in I_1} K_i, \qquad \hat{K} = \bigcap_{i \in I_1} K_i$

Existe un hiperplano que los separa, esto es existe $a \in X$ *tal que*

$$\forall k \in \bar{K}, \tilde{k} \in \hat{K}, \qquad \langle a, k \rangle \geq 0 \geq \langle a, \tilde{k} \rangle$$

Lema B.2.1 *Los conos convexos,cerrados y de vértice* 0*,* $C_1, \ldots, C_s$ *tienen la* **propiedad de separación***, si y solo si existen elementos* $a_1 \in C_1^\circ, \ldots, a_s \in C_s^\circ$ *en los conos duales, al menos uno de ellos diferente de* 0 *tales que:*

$$a_1 + \ldots + a_s = 0.$$

Teorema de Intersección El Teorema de Intersección, es muy importante para demostrar la condiciones necesarias de optimalidad de en la teoría de Boltyanskii.

Teorema B.2.1 (Teorema de Intersección) *Sean* $\Omega_1, \ldots, \Omega_k \subset X$ *y* $x_0 \in \bigcap_{i=1}^k \Omega_i$*,* $C_1, \ldots, C_k$ *cúpulas en el punto* x_0*. Tales que no poseen la propiedad de separación, entonces*

Existe $x \neq x_0$ *tal que:* $\qquad x \in \bigcap_{i=1}^k \Omega_i$

Con este teorema se puede probar el siguiente teorema que es el primer resultado de la Teoría de Optimización de Boltyanskii.

Teorema B.2.2 (Condición Necesaria de Optimalidad General) *Consideremos* $F^0 : A \to \mathbb{R}$*, con* $A \subset X$ *abierto, y con* $F^0 \in C^1(A, \mathbb{R})$ *y los subconjuntos* $\Omega_1, \cdots, \Omega_l \subset X$ *tales que:*

$$\Sigma = \bigcap_{i=1}^{l} \Omega_i \subseteq A.$$

Definamos el problema

$$\min_{z \in \Sigma} F^0 \qquad (\mathcal{P})$$

Entonces si $z_0 \in \Sigma$ *es una solución óptima de* $\mathcal{P}$ *entonces,existen* $\mu \leq 0$ *y* $a_i \in C_i^\circ$ *para* $i \in \{1, \ldots, l\}$ *tales que satisafacen:*

(i) $\mu \leq 0$

(ii) $\mu \cdot DF^0(x_0) = \Sigma_{i=1}^{l} a_i$

(iii) *Si* $\mu = 0$ *entonces* $a_i \neq x_0$

Atendiendo a cada caso particular, se deducen los siguientes teoremas.

Teorema B.2.3 (Restricciones de desigualdad) *Sea* $F^0 : A \to \mathbb{R}$, *con* $A \subset X$ *abierto, y con* $F^0 \in C^1(A, \mathbb{R})$

Para las funciones $F^i : A \to \mathbb{R}$, *con* $i \in \{1, \cdots, k\}$ *y* $F^i \in C^1(A, \mathbb{R})$, *definamos el conjunto:* $\Omega' := \{z \in X : F^i(z) = 0\}$.

Para las funciones $f^j : A \to \mathbb{R}$, *con* $j \in \{1, \cdots, q\}$ *y* $f^j \in C^1(A, \mathbb{R})$, *definamos el conjunto:* $\Omega'' := \{z \in X : f^j(z) \leq 0\}$.

Y con los conjuntos,$\Omega_1, \cdots, \Omega_l \subset X$ *definamos* $\Sigma = \bigcap_{i=1}^{l} \Omega_i$.

Supongamos que: $\Omega' \cap \Omega'' \cap \Sigma \subset A$

Consideremos el problema

$$\begin{array}{cc} \textit{Minimizar} & F^0 \\ z \in \Sigma \cap \Omega' \cap \Omega'' & \end{array} \qquad (\mathcal{P})$$

Entonces si $z_0 \in \Sigma \cap \Omega' \cap \Omega''$ *una solución del problema. Entonces existen* $\mu_0, \mu_1, \ldots, \mu_k \in \mathbb{R}$, $\lambda_j \geq 0 \quad j \in I(z_0)$,*y elementos* $a_i \in M_i^\circ, \quad i = \{1, \ldots, l\}$ *tales que:*

(α) $\mu_0 \leq 0$

(β) $\sum_{i=0}^{k} \mu_i DF^i(z_0) + \sum_{j \in I(z_0)} \lambda_j Df^j(z_0) = a_1 + \ldots + a_l$

(γ) *Si* $\mu_0 = \ldots = \mu_k = 0$ *y* $\lambda_j = 0, \quad j \in I(z_0)$, *entonces algún* a_i *será distinto de cero.*

Donde $M_1, \ldots, M_l$ *cúpulas de* x_0 *para cada* Ω_i, *respectivamente. Y* $I(z_0)$ *las restricciones activas para* Ω'.

Proposición B.2.1 *Sea* $F^0 : A \to \mathbb{R}$, *con* $A \subset X$ *abierto, y con* $F^0 \in C^1(A, \mathbb{R})$ *sean los conjuntos* $\Omega_1, \ldots, \Omega_n \subset X$ *y también* $Z_1, \ldots, Z_m \subset X$

Definamos $\Sigma := \Omega_1 \cap \ldots \cap \Omega_l \cap Z_1 \cap \ldots \cap Z_m$

Tenemos el siguiente problema:

$$\min_{z \in \Sigma} F^0 \qquad (\mathcal{P})$$

Sea z_0 *una solución optimal. Considerando* $M_1, \ldots, M_n$ *y* $N_1, \ldots, N_m$ *cúpulas en* z_0 *de* Ω_i *y* Z_j *respectivamente. Supongamos además que* $N_1, \ldots, N_m$ *no poseen la propiedad de separación.*

Entonces, existen $\mu_0 \in \mathbb{R}$ *y* $a_i \in M_i^\circ$ *que satisfacen:*

(i) $\mu_0 \leq 0$

(ii) $\forall \delta z \in N_1 \cap \ldots \cap N_m \qquad \langle \mu_0 DF^0(z_0) - a_1 - \ldots - a_n ; \delta z \rangle \leq 0$

(iii) Si $\mu_0 = 0$ *encontes algún* $a_i \neq z_0$

De la proposición anterior se desprende el siguiente resultado que será utilizado para demostrar el Principio de Pontryaguin desde la aproximación de Boltyanskii.

En primer lugar plantémos el siguiente problema de optimización

$$\max_{\substack{F^i(w)=0,\\ w\in\Xi}} F^0(w), \tag{$\mathcal{P}$}$$

donde para $i = 0, \cdots, N$, las funciones $F^i : A \to \mathbb{R}$ están definidas en $A \subset X$ (abierto), y son de tipo C^1.

Teorema B.2.4 (Condición Necesaria de Optimalidad de Boltyanskii) *Consideremos* z_0 *una solución óptima del problema* $(\mathcal{P})$.

Entonces existen números $\mu_0, \mu_1, \cdots, \mu_k \in \mathbb{R}$, *tales que:*

1. $\mu_0 \leq 0$
2. *Para todo* $\delta w \in L$
$$\left\langle \sum_{i=0}^{k} \mu_i DF^i(z_0) \ ; \ \delta w \right\rangle \leq 0,$$
donde L *es la cúpula de* z_0 *respecto a* Ξ.

Bibliografía

[1] M.S. Bazaraa, H.D. Sherali y C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, John Wiley and Sons, New Jersey, 2006.

[2] R. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, New Jersey, (1957).

[3] R. Bellman y S. Dreyfus, *Applied Dynamic Programmig*, Princeton University Press, Princeton, New Jersey, (1962).

[4] A. Bensoussan, E.G. Hurst y B. Naslund, *Management Applications of Modern Control Theory*, Nort-Holland, 1974.

[5] D. Bertsekas, *Dynamic Programming and Optimal Control, volume 1 and 2*, Athena Scientific, Belmont, Massachusetts, second edition, 2000.

[6] J. Blot y H. Chebbi, *Discrete Time Pontryaguin Principles with Infinite Horizon*, CERMSEN, M.S.E, Université de Paris I, Paris, 2000.

[7] V.G. Boltyanskii, *Commande Optimale des Systèmes Discrets*, MIR, Moscú, 1976.

[8] E. Cerdá, *Optimización Dinámica*, Prentice Hall, Madrid, 2001.

[9] C.W. Clark, *Mathematical Bioeconomics*, John Wiley and Sons, New York, second edition, 1990.

[10] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983.

[11] P. Dasgupta, *The Control of Resources*, Basil Blackwell, Oxford, 1982.

[12] M. De Lara y L. Doyen, *Sustainable Management of Natural Resource: Mathematical Models and Methods*, Springer-Verlag, Berlin, 2008.

[13] L.T. Fan y C.S. Wang, *Décisions économiques séquentielle optimales*, Dunod, Paris, 1971.

[14] T. Friesz, *Dynamic Optimization and Differential Games*, International Series in Operations Research and Management Science Vol. 135, Springer Verlag, 2010.

[15] A. Izmailov y M. V. Solodov, *Otimizacão Vol.1*, IMPA, Rio de Janeiro, 2005.

[16] E. Lages, *Algebra Lineal*, IMCA, Lima, 1998.

[17] M.M. Mäkelä, P. Neittaanmäki, *Nonsmooth optimization: analysis and algorithms with applications to optimal control*, World Scientific, Singapore, 1992.

[18] L.W. McKenzie, *Optimal Economic Growth, Turnpike Theorems and Comparative Dynamics, Chapter 26 in "Handbook of Mathematical Economics"*, Vol. III, K.J. Arrow and M.D. Intriligator (Editors), Elsevier Sc. Publ. B.V. (North-Holland), 1986.

[19] P. Michel, *Programmes Mathématiques Mixtes. Application au principe du maximum en temps discret dans le cas déterministe et dans le cas stochastique* , R.A.I.R.O. **14** (1980), 1-19.

[20] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, vol. 4. Interscience, 1962.

[21] T.J. Sargent, *Dynamic Macroeconomic Theory*, Harvard University Press, Cambridge, Massachusetts, 1987.

[22] N.L. Stokey, R.E. Lucas y E.C. Prescott, *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, Massachusetts, 1989.

[23] W.L. Winston y J.B. Goldberg, *Operations research: applications and algorithms*, Thomson Brooks Cole, 2004.