UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

NORMALIDAD Y UNICIDAD DE MULTIPLICADORES DE LAGRANGE

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTORA EN CIENCIAS

PRESENTA:
KARLA LORENA CORTEZ DEL RÍO

TUTOR PRINCIPAL
JAVIER FERNANDO ROSENBLUETH LAGUETTE
UNAM-IIMAS

COMITÉ TUTOR
PABLO BARBERIS BLOSTEIN
UNAM-IIMAS

MARÍA DE LA LUZ JIMENA DE TERESA DE OTEYZA
UNAM-INSTITUTO DE MATEMÁTICAS

CIUDAD UNIVERSITARIA CD. MX. MAYO 2018.

Agradecimientos

Por la culminación de esta tesis quiero agradecer a mi familia, especialmente a mis sobrinos Gael y Lorena, por darme tanta alegría y motivación. A todos y cada uno de mis maestros de licenciatura y maestría, por todo lo que aprendí de ellos y todo el apoyo que me brindaron. A Javier, por estos años en los que me ha proporcionado su apoyo, su confianza y, sobre todo, su paciencia. A Jorge, por compartir tantas horas de trabajo y ayudarme siempre que las ideas escaseaban. Y a Dios, por ponerlos a todos en mi camino y hacer posible este trabajo.

Agradezco también a todos los sinodales, de manera particular al Dr. Héctor Sánchez Morgado, por el tiempo que dedicaron a la revisión y mejoramiento del trabajo.

Introducción

Para los problemas de optimización con restricciones en espacios de dimensión finita, las cuestiones referentes a condiciones de primer orden, así como las tres preguntas fundamentales que surgen al tratar con condiciones necesarias de segundo orden (es decir, qué funciones las satisfacen, bajo qué hipótesis y en qué conjuntos tienen lugar), han sido respondidas de manera satisfactoria en términos de multiplicadores de Lagrange. Los resultados que responden estas interrogantes se pueden encontrar en textos clásicos como [23] o [25] donde (utilizando las nociones de normalidad y regularidad) se muestra que, dado un multiplicador de Lagrange λ, la normalidad de una solución con respecto a un conjunto (que denotaremos como S_1) definido por el signo de las componentes de λ garantiza la no negatividad de cierta forma cuadrática en el cono de restricciones tangenciales asociado a la solución y al conjunto S_1.

Esta condición de normalidad es llamada, por algunos autores, "calificación de restricción estricta de Mangasarian-Fromovitz" (ver [21, 26]) la cual, como muestra Kyparisis en [26], resulta ser necesaria y suficiente para la unicidad del multiplicador λ. Sin embargo, como se explica en [23, pág. 310], no es apropiado llamar calificación de restricción a dicha condición pues ésta requiere de la existencia de un multiplicador de Lagrange conocido de antemano. En [44], Wachsmuth proporciona un argumento similar y prueba que la calificación de restricción más débil que garantiza la existencia y unicidad de multiplicadores de Lagrange es la condición de independencia lineal.

La unicidad de multiplicadores de Lagrange asociados a una solución es un tema de innegable importancia tanto desde el punto de vista teórico como práctico, pues éstos se emplean para resolver una gran diversidad de problemas en áreas como ingeniería, física y economía, por mencionar sólo algunas. Dentro de las matemáticas, el problema de cuándo es posible garantizar que un multiplicador de Lagrange exista y sea único ha sido estudiado en varios problemas de optimización sujeta a restricciones, como el análisis de sensibilidad (ver [27, 28]), la optimización multiobjetivo

[11], el análisis de convergencia de algoritmos de optimización [12], la optimización sujeta a restricciones de cono [42], o la optimización compuesta [20]. De hecho, en esta última referencia, se muestra que la unicidad del multiplicador de Lagrange está caracterizada por una condición que extiende de manera natural a la calificación de restricción estricta de Mangasarian-Fromovitz.

Otra es la situación en áreas como cálculo de variaciones y control óptimo, donde la teoría de unicidad de multiplicadores de Lagrange ha recibido menor atención pues el tema apenas se menciona en algunos textos sobre condiciones de segundo orden como [19] y [24]. Uno de los escasos trabajos donde se analiza más detalladamente es el de Malanowski, quien deriva en [32] algunas condiciones de suficiencia para la existencia de un único multiplicador de Lagrange normal en un problema de control óptimo con restricciones de estado y mixtas control-estado. Entre estas condiciones están la independencia lineal puntual de los gradientes de restricciones α-activas y la controlabilidad de la ecuación de estado linealizada.

La contribución más importante del presente trabajo a esta área es dar condiciones necesarias y suficientes para la unicidad de multiplicadores de Lagrange (satisfaciendo condiciones necesarias de primer orden del problema de Lagrange con puntos extremos fijos), en los contextos de cálculo de variaciones y control óptimo sujeto a restricciones en forma de igualdad y desigualdad tanto isoperimétricas como puntuales (ver [7, 17]) .

Comenzaremos con un resumen de los principales aspectos que relacionan las nociones de normalidad, unicidad de los multiplicadores de Lagrange y condiciones necesarias de segundo orden en el caso de dimensión finita, lo que permitirá una mejor comprensión de la teoría que desarrollaremos en los capítulos posteriores. En el segundo capítulo analizamos el problema isoperimétrico de Lagrange, tanto en el contexto de cálculo de variaciones como de control óptimo, obteniendo para ambos casos resultados análogos a los de Kyparisis y Wachsmuth, caracterizando de esta manera la unicidad de los multiplicadores de Lagrange. En el mismo capítulo, damos un breve resumen de los resultados concernientes a condiciones de segundo orden obtenidos en [5] pues, como veremos, éstos nos permitirán concluir que la unicidad garantiza condiciones de segundo orden.

A lo largo del tercer capítulo, abordaremos el problema de Lagrange en control óptimo con restricciones de igualdad y desigualdad en los controles, problema que, como explica Clarke en [14, pág. 335], resulta de mayor complejidad que los anteriores debido, en parte, a que ahora tenemos una infinidad de restricciones, una por cada punto del intervalo considerado. Mostraremos las similitudes y diferencias entre éste y los casos anteriores, en relación con los conceptos de normalidad, regularidad y condiciones del tipo Mangasarian-Fromovitz, así como las implicaciones que éstas tienen en los multiplicadores de Lagrange y las condiciones de segundo orden. En particular veremos que, para este tipo de problemas en control óptimo, la unicidad de los multiplicadores de Lagrange, sorprendentemente, no está caracterizada por la normalidad con respecto a S_1. En este trabajo logramos obtener la caracterización deseada a través de una nueva condición que es implicada por la anterior. Terminamos el capítulo con otro resultado teórico de gran relevancia el cual, sin duda, cambia radicalmente la perspectiva sobre el tema. Probamos que el vínculo existente en programación no lineal (ver Teorema 1.2.9) que relaciona los conceptos de normalidad y regularidad, deja de ser válido para problemas de cálculo de variaciones que involucran restricciones en forma de igualdades y desigualdades.

Concluimos este trabajo con un resumen de los principales resultados obtenidos y explicamos cómo extenderlos a contextos más generales, así como posibles líneas de investigación futura.

Contenido

Capítulo 1

Optimización en espacios de dimensión finita

En este capítulo abordaremos el problema de optimización sujeta a restricciones en espacios de dimensión finita, conocido también como programación no lineal, al cual denotaremos como $N(S)$. Este problema consiste en minimizar una función real-valuada f en un subconjunto S de $\mathbb{R}^n$. Nuestro objetivo en este apartado es utilizar los conceptos de normalidad y regularidad para derivar, de forma clara y sencilla, condiciones necesarias para optimalidad de primer y segundo orden cuando el subconjunto S está definido por restricciones en forma de igualdad y desigualdad. Los resultados (cuyos detalles se pueden consultar en [25]) se expresan en términos de los multiplicadores de Karush-Kuhn-Tucker, más comúnmente conocidos como multiplicadores de Lagrange.

Como veremos más adelante, la importancia de la regularidad radica principalmente en que garantiza condiciones de primer orden, es decir, si x_0 es un punto regular que minimiza a f localmente en S, entonces existen multiplicadores de Lagrange asociados a dicha solución. Sin embargo, determinar si un punto es regular puede no resultar sencillo utilizando únicamente la definición. Por ello, se buscaron criterios que permitieran asegurar esta condición y se verificaran más fácilmente; de ellos, el más común es el de normalidad.

Por otra parte, para tener condiciones de segundo orden, en general se requieren hipótesis más fuertes pues, como mostraremos por medio de algunos ejemplos, la normalidad no es suficiente para garantizarlas. Con este fin, se introdujeron nuevas versiones (más restrictivas) de esta condición que permiten garantizar no sólo la existencia sino también la unicidad de multiplicadores de Lagrange, así como la validez de condiciones de segundo orden en ciertos subconjuntos de $\mathbb{R}^n$ asociados a una solución x_0 que poseen estructura de cono convexo.

Comenzaremos analizando el caso donde S es un subconjunto arbitrario de $\mathbb{R}^n$ y daremos condiciones necesarias de primer y segundo orden en términos del cono tangente a S en un punto $x_0 \in S$. La noción de cono tangente que damos a continuación se debe a Hestenes [25]. Sin embargo, como se muestra en [23], esta noción es equivalente a la introducida en 1932 por Bouligand, quien lo llamó cono contingente a S en x_0. También en [23] podemos encontrar múltiples definiciones equivalentes de autores como Bazaraa, Goode, Nashed, Kurcyusz, Rockafellar, Saks, Rogak, entre otros.

En lo sucesivo, '$*$' denotará transpuesta y supondremos que f tiene diferencial en $x_0 \in S$ o segunda diferencial cuando aparezcan segundas derivadas.

Definición 1.0.1. *Sean x_0 un punto y h un vector unitario en $\mathbb{R}^n$. Decimos que la sucesión $\{x_q\} \subset \mathbb{R}^n$ converge a x_0 en la dirección h, si $x_q \neq x_0$ y*

$$\lim_{q\to\infty} |x_q - x_0| = 0, \qquad \lim_{q\to\infty} \frac{x_q - x_0}{|x_q - x_0|} = h.$$

El cono tangente a S en x_0, denotado como $T_S(x_0)$, es el cono cerrado generado por los vectores unitarios h para los cuales existe una sucesión $\{x_q\} \subset S$ que converge a x_0 en la dirección h.

Equivalentemente (ver [25]), $T_S(x_0)$ es el cono compuesto por todos los vectores $h \in \mathbb{R}^n$ para los cuales existen sucesiones $\{x_q\} \subset S$ y $\{t_q\} \subset \mathbb{R}^+$, tales que

$$\lim_{q\to\infty} t_q = 0, \qquad \lim_{q\to\infty} \frac{x_q - x_0}{t_q} = h.$$

Observemos que, si f tiene diferencial en x_0 y $\{x_q\}$ es una sucesión en S que converge a x_0 en la dirección h, entonces

$$\lim_{q\to\infty} \frac{f(x_q) - f(x_0)}{|x_q - x_0|} = f'(x_0, h).$$

Si además f tiene segunda diferencial, entonces

$$\lim_{q\to\infty} \frac{f(x_q) - f(x_0) - f'(x_0, x_q - x_0)}{|x_q - x_0|^2} = \frac{1}{2} f''(x_0, h).$$

Utilizando estas igualdades, podemos probar fácilmente el siguiente resultado.

Teorema 1.0.2. *Supongamos que $x_0 \in S$ resuelve $N(S)$ localmente, es decir, x_0 minimiza a f localmente en S. Entonces $f'(x_0, h) \geq 0$ para todo $h \in T_S(x_0)$. Más aún, si $f'(x_0) = 0$, entonces $f''(x_0, h) \geq 0$ para todo $h \in T_S(x_0)$.*

Demostración. Sea $h \in T_S(x_0)$ un vector unitario y $\{x_q\}$ una sucesión en S que converge a x_0 en la dirección h. Como x_0 es un punto de mínimo local, para valores suficientemente grandes de q se debe cumplir $f(x_q) \geq f(x_0)$. Por lo tanto,

$$0 \leq \lim_{q \to \infty} \frac{f(x_q) - f(x_0)}{|x_q - x_0|} = f'(x_0, h).$$

Si $f'(x_0) = 0$, entonces

$$0 \leq \lim_{q \to \infty} \frac{f(x_q) - f(x_0)}{|x_q - x_0|^2} = \frac{1}{2} f''(x_0, h). \qquad \square$$

1.1. Restricciones de igualdad

En esta sección consideramos el caso donde $N(S)$ consiste en minimizar f en el conjunto S determinado por restricciones de igualdad, es decir,

$$S := \{x \in \mathbb{R}^n \mid g_\alpha(x) = 0 \ (\alpha \in A)\},$$

donde $A = \{1, \ldots, m\}$ $(m \leq n)$ y $f, g_1, \ldots, g_m$ son funciones de $\mathbb{R}^n$ a $\mathbb{R}$, continuas en una vecindad de $x_0 \in S$ y con segunda diferencial en x_0.

Definición 1.1.1. *El conjunto de vectores que satisfacen las restricciones tangenciales en $x_0 \in S$ se define como*

$$R_S(x_0) := \{h \in \mathbb{R}^n \mid g'_\alpha(x_0, h) = 0 \ (\alpha \in A)\}.$$

Con un razonamiento análogo al de la prueba del Teorema 1.0.2, es fácil verificar que $T_S(x_0) \subset R_S(x_0)$. Diremos que x_0 es un punto regular con respecto a S si $T_S(x_0) = R_S(x_0)$.

La condición de regularidad con respecto a S es importante porque nos da una buena caracterización del cono tangente $T_S(x_0)$ puesto que, con la definición original (Definición 1.0.1), podría resultar difícil determinar la pertenencia de un vector h a dicho cono mientras que, en general, es más sencillo verificar que satisfaga las restricciones tangenciales. Más aún, como mencionamos en la introducción, esta condición permite asegurar la condición de primer orden conocida como *regla de los multiplicadores de Lagrange*, la cual es consecuencia del Teorema 1.0.2 y el siguiente resultado auxiliar cuya demostración puede consultarse en [24].

Lema 1.1.2. *Sean L, L_i $(i \in A := \{1, \ldots, m\})$ funcionales lineales en un espacio vectorial real X, defina*

$$R := \{x \in X \mid L_i(x) = 0 \ (i \in A)\},$$

y suponga que $L(x) = 0$ para todo $x \in R$. Entonces existe $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ tal que $L(x) = \sum_1^m \lambda_i L_i(x)$ $(x \in X)$. Si $\{L_i\}_1^m$ son linealmente independientes, entonces λ es único.

Teorema 1.1.3. Regla de los multiplicadores de Lagrange. *Suponga que x_0 resuelve $N(S)$ localmente y es un punto regular con respecto a S. Entonces existe $\lambda \in \mathbb{R}^m$ tal que, si $F(x) := f(x) + \langle \lambda, g(x) \rangle$, se tiene que $F'(x_0) = 0$.*

Demostración. Por el Teorema 1.0.2 y la regularidad con respecto a S de x_0, $f'(x_0, h) \geq 0$ para todo $h \in R_S(x_0)$, pero éste es un subespacio vectorial de $\mathbb{R}^n$ por lo que $-h \in R_S(x_0)$ y, por tanto, $f'(x_0, h) = 0$ para todo $h \in R_S(x_0)$. La conclusión se sigue entonces por el Lema 1.1.2. □

El siguiente resultado también es consecuencia directa del Teorema 1.0.2. Para probarlo, basta observar que $F(x) = f(x)$ para $x \in S$.

Teorema 1.1.4. *Suponga que existe $\lambda \in \mathbb{R}^m$ tal que $F'(x_0) = 0$ donde $F(x) := f(x) + \langle \lambda, g(x) \rangle$. Si x_0 es una solución local de $N(S)$, entonces $F''(x_0, h) \geq 0$ para todo $h \in T_S(x_0)$. En particular, la conclusión es cierta para $h \in R_S(x_0)$ si x_0 es regular con respecto a S.*

Como también mencionamos ya, un criterio fácil de verificar que permite determinar si un punto es regular es la condición de normalidad, que definimos a continuación.

Definición 1.1.5. *Decimos que x_0 es un punto normal con respecto a S si los gradientes $g_1'(x_0), \ldots, g_m'(x_0)$ son linealmente independientes, es decir, si la matriz*

$$g'(x_0) := \left(\frac{\partial g_\alpha(x_0)}{\partial x^i} \right) \quad (\alpha = 1, \ldots, m; \ i = 1, \ldots, n)$$

es de rango m.

Teorema 1.1.6. *Si x_0 es un punto normal con respecto a S, entonces x_0 es un punto regular con respecto a S.*

Si g es C^1 en una vecindad de x_0, se puede dar una prueba sencilla del Teorema 1.1.6 utilizando un subconjunto de $T_S(x_0)$ conocido como cono de vectores tangentes curvilíneos, y el Lema 1.1.8, el cual es consecuencia del teorema de la función implícita. Daremos una demostración más general en la próxima sección, donde S está determinado por restricciones de igualdad y desigualdad.

Definición 1.1.7. *El conjunto de vectores tangentes curvilíneos a S en x_0 se define como*

$$C_S(x_0) := \{h \in \mathbb{R}^n \mid \exists \delta > 0 \text{ y } x \colon [-\delta, \delta] \to S \text{ tal que } x(0) = x_0 \text{ y } \dot{x}(0) = h\}.$$

Lema 1.1.8. *Supongamos que $g \colon \mathbb{R}^n \to \mathbb{R}^m$ $(m \leq n)$ es de clase C^k $(k \geq 1)$ en una vecindad de un punto x_0 y que $g'(x_0)$ es de rango completo. Entonces, dado $h \in \mathbb{R}^n$, existen $\delta > 0$ y $x \colon [-\delta, \delta] \to \mathbb{R}^n$ de clase C^k tales que*

$$x(0) = x_0, \quad \dot{x}(0) = h \quad \text{y} \quad g(x(t)) = g(x_0) + tg'(x_0, h) \quad (|t| \leq \delta).$$

Demostración. Sea $h \in \mathbb{R}^n$ y definamos, para $(t, b) \in \mathbb{R} \times \mathbb{R}^m$,

$$\bar{x}(t, b) := x_0 + th + g'^*(x_0)b \quad \text{y} \quad G(t, b) := g(\bar{x}(t, b)) - g(x_0) - tg'(x_0, h).$$

Como $G(0, 0) = 0$ y $|G_b(0, 0)| = |g'(x_0)g'^*(x_0)| \neq 0$, por el teorema de la función implícita existen $\delta > 0$ y $b \colon [-\delta, \delta] \to \mathbb{R}^m$ de clase C^k tales que

$$b(0) = 0 \quad \text{y} \quad G(t, b(t)) = 0 \quad (t \in [-\delta, \delta]).$$

Derivando con respecto a t y evaluando en $t = 0$, obtenemos

$$0 = G_t(0, 0) + G_b(0, 0)b'(0) = G_b(0, 0)b'(0)$$

y, como $|G_b(0, 0)| \neq 0$, se debe cumplir que $b'(0) = 0$. Sea $x(t) := \bar{x}(t, b(t))$ para todo $t \in [-\delta, \delta]$. Por lo tanto, $x(0) = x_0$, $\dot{x}(0) = \bar{x}_t(0, 0) + \bar{x}_b(0, 0)b'(0) = h$ y, para

$|t| \leq \delta$, se tiene

$$g(x(t)) - g(x_0) - tg'(x_0, h) = G(t, b(t)) = 0. \qquad \square$$

Demostración del Teorema 1.1.6. Sea $h \in R_S(x_0)$. Por el Lema 1.1.8, existen $\delta > 0$ y $x\colon [-\delta, \delta] \to \mathbb{R}^n$ de clase C^1 tales que $x(0) = x_0$, $\dot{x}(0) = h$ y, para todo $t \in [-\delta, \delta]$ y $\alpha \in A$, $g_\alpha(x(t)) = tg'_\alpha(x_0, h) = 0$. Por tanto, $x(t) \in S$ para t en dicho intervalo. Esto implica que $h \in C_S(x_0) \subset T_S(x_0)$. $\square$

1.2. Restricciones de igualdad y desigualdad

En esta sección analizaremos el caso donde $S \subset \mathbb{R}^n$ está determinado por un conjunto finito de restricciones en forma de igualdad y desigualdad. Consideremos las funciones $f, g_i\colon \mathbb{R}^n \to \mathbb{R}$ $(i \in A \cup B)$ y definamos el conjunto

$$S = \{x \in \mathbb{R}^n \mid g_\alpha(x) \leq 0 \ (\alpha \in A),\ g_\beta(x) = 0 \ (\beta \in B)\},$$

donde $A = \{1, \ldots, p\}$, $B = \{p+1, \ldots, m\}$, (en general puede ocurrir que $m > n$). Los casos $p = 0$ y $p = m$ tienen las interpretaciones obvias. Supongamos también que f y g_i son continuas en una vecindad de $x_0 \in S$ y poseen segunda diferencial en x_0.

Definición 1.2.1. *Definimos el conjunto de índices activos en x_0 por*

$$I(x_0) := \{\alpha \in A \mid g_\alpha(x_0) = 0\}$$

y el cono linealizado de S en x_0 como el conjunto de vectores que satisfacen las restricciones tangenciales en x_0, es decir,

$$R_S(x_0) := \{h \in \mathbb{R}^n \mid g'_\alpha(x_0, h) \leq 0 \ (\alpha \in I(x_0)),\ g'_\beta(x_0, h) = 0 \ (\beta \in B)\}.$$

Notemos que $T_S(x_0) \subset R_S(x_0)$ y, por tanto, la contención se invierte cuando se consideran los conos polares, es decir, $R_S^*(x_0) \subset T_S^*(x_0)$[1]. Diremos que x_0 es un punto regular con respecto a S si $T_S(x_0) = R_S(x_0)$.

En general, se conocen como "calificaciones de restricción" a las condiciones im-

[1]Dado $X \subset \mathbb{R}^n$, $X^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \leq 0$ para todo $x \in X\}$.

puestas sobre las restricciones g_i, que son independientes de la estructura geométrica de S y permiten asegurar que $R_S^*(x_0)$ coincide con $T_S^*(x_0)$ pues, utilizando teoría de conos convexos como en [25], o el lema de Farkas como en [23], se puede mostrar que $R_S^*(x_0) = T_S^*(x_0)$ es la hipótesis más débil con la que se obtiene la condición de primer orden de los multiplicadores de Lagrange. Claramente, la regularidad con respecto a S es una de estas condiciones, y es conocida también como "calificación de restricción de Abadie."

Como ocurre en el caso con restricciones de igualdad, bajo la hipótesis de regularidad (con respecto a S), la existencia de multiplicadores es consecuencia del Teorema 1.0.2 y del siguiente resultado auxiliar sobre funcionales lineales cuya demostración puede encontrarse en [24, 25].

Lema 1.2.2. *Sean L, L_i ($i \in A \cup B,\ A = \{1, \ldots, p\},\ B = \{p+1, \ldots, m\}$) funcionales lineales en el espacio vectorial real X y definamos*

$$R = \{x \in X \mid L_\alpha(x) \leq 0 \ (\alpha \in A),\ L_\beta(x) = 0 \ (\beta \in B)\}.$$

Si $L(x) \geq 0$ para toda $x \in R$, entonces existe $\lambda \in \mathbb{R}^m$ tal que

$$\lambda_\alpha \geq 0 \quad (\alpha \in A) \quad y \quad L(x) + \sum_1^m \lambda_i L_i(x) = 0 \quad (x \in X).$$

Si $\{L_i\}_1^m$ es linealmente independiente, entonces λ es único.

Teorema 1.2.3. Regla de los multiplicadores de Lagrange de primer orden. *Supongamos que x_0 resuelve $N(S)$ localmente. Si x_0 es un punto regular con respecto S entonces existe $\lambda \in \mathbb{R}^m$ tal que*

i. *$\lambda_\alpha \geq 0$ y $\lambda_\alpha g_\alpha(x_0) = 0$ ($\alpha \in A$).*
ii. *Si $F(x) := f(x) + \langle \lambda, g(x) \rangle$ entonces $F'(x_0) = 0$.*

Demostración. Por el Teorema 1.0.2 y la regularidad con respecto a S de x_0, $f'(x_0; h) \geq 0$ para toda $h \in R_S(x_0)$. La conclusión se sigue por el Lema 1.2.2. □

Como en el caso de condiciones de primer orden, la regla de los multiplicadores de Lagrange de segundo orden se sigue aplicando el Teorema 1.0.2.

Teorema 1.2.4. Regla de los multiplicadores de Lagrange de segundo orden. *Supongamos que existe $\lambda \in \mathbb{R}^n$ tal que*

i. *$\lambda_\alpha \geq 0$ y $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.*

ii. *Si $F(x) := f(x) + \langle \lambda, g(x) \rangle$ entonces $F'(x_0) = 0$.*

Si x_0 es una solución local de $N(S)$ entonces $F''(x_0, h) \geq 0$ para toda $h \in T_{S_1}(x_0)$ donde $S_1 := S_1(\lambda) = \{x \in S \mid F(x) = f(x)\}$.

Demostración. Como x_0 minimiza F en S_1 y $F'(x_0) = 0$, el resultado se sigue por Teorema 1.0.2. □

Es importante tener claro que el conjunto S_1 definido arriba depende de los multiplicadores de Lagrange $\lambda_i \in \mathbb{R}$ $(i = 1, \ldots, m)$, y la condición de segundo orden es válida en el cono tangente de S_1 en x_0 y no necesariamente en el cono tangente de S en x_0, como ocurre en el caso con restricciones de igualdad.

Si $\lambda \in \mathbb{R}^m$ satisface (i) y (ii) del Teorema 1.2.3 y $\Gamma := \{\alpha \in A \mid \lambda_\alpha > 0\}$, entonces

$$\begin{aligned} S_1 &= \{x \in \mathbb{R}^n \mid g_\alpha(x) \leq 0 \ (\alpha \in A,\ \lambda_\alpha = 0),\ g_\beta(x) = 0 \ (\beta \in \Gamma \cup B)\} \\ &= \{x \in S \mid g_\alpha(x) = 0 \ (\alpha \in \Gamma)\}. \end{aligned}$$

Por lo tanto, el cono linealizado de S_1 en x_0 es

$$\begin{aligned} R_{S_1}(x_0) &= \{h \in \mathbb{R}^n \mid g'_\alpha(x_0, h) \leq 0 \ (\alpha \in I(x_0),\ \lambda_\alpha = 0), \\ &\qquad\qquad\qquad\qquad g'_\beta(x_0, h) = 0 \ (\beta \in \Gamma \cup B)\} \\ &= \{h \in R_S(x_0) \mid g'_\alpha(x_0, h) = 0 \ (\alpha \in \Gamma)\} \\ &= \{h \in R_S(x_0) \mid f'(x_0, h) = 0\}. \end{aligned}$$

Observemos que, si en el Teorema 1.2.4, x_0 es un punto regular con respecto a S_1, entonces $F''(x_0, h) \geq 0$ para todo $h \in R_{S_1}(x_0)$.

Para este problema, el concepto de normalidad con respecto a S puede ser controversial. En particular (por poner un ejemplo), Hestenes da en [24] y [25] definiciones que no son equivalentes. En [24], la definición se basa en la condición de normalidad del caso anterior donde S está determinado únicamente por igualdades pues hace notar que, sin perder generalidad, podemos suponer que todos los índices son activos en x_0 ya que, si para algún $i \in A$ se tiene que $g_i(x_0) < 0$, entonces $g_i(x) < 0$ en una vecindad de x_0. Consecuentemente, en una vecindad suficientemente pequeña de x_0, el conjunto S está completamente determinado por las igualdades y desigualdades con α en el conjunto de índices para los cuales $g_\alpha(x_0) = 0$. Más aún, si $I(x_0) \neq A$, los resultados que se obtienen son válidos si nos restringimos a los índices activos.

Con base en la observación anterior, la condición de normalidad con respecto a S se define como sigue.

Definición 1.2.5. *Diremos que un punto x_0 es s-normal con respecto a S si los gradientes $g_i'(x_0)$ $(i \in I(x_0) \cup B)$ son linealmente independientes.*

Teorema 1.2.6. *Si x_0 es s-normal con respecto a S entonces x_0 es un punto regular con respecto a S.*

Demostración. La prueba es similar a la del Teorema 1.1.6. Definamos al conjunto de vectores *tangentes curvilíneos positivos* de S en x_0 por

$$C_S^+(x_0) := \{h \in \mathbb{R}^n \mid \exists \delta > 0 \text{ y } x\colon [0,\delta) \to S \text{ tales que } x(0) = x_0 \text{ y } \dot{x}(0) = h\}.$$

Notemos que $C_S^+(x_0) \subset T_S(x_0) \subset R_S(x_0)$. Supongamos que g es C^1 (en otro caso, ver [24]). Por la observación previa, es suficiente mostrar que $R_S(x_0) \subset C_S^+(x_0)$. Sea $h \in R_S(x_0)$. Por el Lema 1.1.8, $\exists \delta > 0$ y $x\colon [-\delta,\delta] \to \mathbb{R}^n$ de clase C^1 tales que $x(0) = x_0$, $\dot{x}(0) = h$ y, para todo $t \in [-\delta,\delta]$ y $\alpha \in I(x_0) \cup B$, $g_\alpha(x(t)) = tg_\alpha'(x_0,h)$. Más aún, si es necesario, es posible disminuir δ de modo que, para $\alpha \notin I(x_0)$, se satisfaga $g_\alpha(x(t)) < 0$ en $[0,\delta)$. Así, $x(t) \in S$ para $t \in [0,\delta]$. Esto prueba que $h \in C_S^+(x_0)$. □

El resultado anterior es precisamente el enunciado del Lema 10.1, Capítulo 1 de [24] y tiene una implicación particular con respecto al conjunto S_1, específicamente, que x_0 sea s-normal con respecto a S implica que es s-normal con respecto a S_1, por tanto, regular con respecto a S_1 y usando el Teorema 1.2.4 probamos el siguiente teorema.

Teorema 1.2.7. *Supongamos que existe $\lambda \in \mathbb{R}^m$ tal que*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.

ii. *Si* $F(x) := f(x) + \langle \lambda, g(x) \rangle$ *entonces* $F'(x_0) = 0$.

Si x_0 resuelve $N(S)$ localmente y es un punto s-normal con respecto a S_1, entonces $F''(x_0,h) \geq 0$ para todo $h \in R_{S_1}(x_0)$.

Por otro lado, en [25], Hestenes define la normalidad con respecto a S de la siguiente manera.

Definición 1.2.8. *Un punto x_0 es normal con respecto a S si $\lambda = 0$ es la única solución del sistema*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.

ii. $\sum_1^m \lambda_i g_i'(x_0) = 0$.

En [25] se prueba que x_0 es normal con respecto a S si y sólo si se cumple lo siguiente:

(MF) *Mangasarian-Fromovitz en* x_0.

El conjunto $\{g'_\beta(x_0) \mid \beta \in B\}$ es linealmente independiente y, si $p > 0$, $\exists h \in \mathbb{R}^n$ tal que

$$g'_\alpha(x_0; h) < 0 \ (\alpha \in I(x_0)), \quad g'_\beta(x_0; h) = 0 \ (\beta \in B).$$

La condición anterior es conocida como la "calificación de restricción de Mangasarian-Fromovitz", y es una caracterización fundamental de la normalidad con respecto a S que permite probar la relación entre normalidad y regularidad.

Teorema 1.2.9. *Si x_0 es un punto normal con respecto a S entonces x_0 es un punto regular con respecto a S.*

Este teorema es el resultado básico que relaciona los conceptos de normalidad y regularidad. La demostración puede consultarse en [23, 25]. La idea central de ésta consiste en utilizar las propiedades del cono tangente $T_S(x_0)$ y la caracterización anterior para mostrar que todo $h \in R_S(x_0)$ es el límite de una sucesión $\{h_n\} \subseteq T_S(x_0)$ y, como este último es cerrado, debe contener a h.

En [34], McShane, basándose en la teoría de aumentabilidad, proporciona una prueba sencilla de la condición de primer orden de los multiplicadores de Lagrange, mostrando que la siguiente regla "extendida" tiene lugar. Esta regla de los multiplicadores extendida, o condición de Fritz John, conduce de manera natural a la Definición 1.2.8, ya que ésta implicaría un multiplicador de costo λ_0 positivo.

Teorema 1.2.10. *Si x_0 resuelve $N(S)$ localmente, entonces existen $\lambda_0 \geq 0$ y $\lambda \in \mathbb{R}^m$ tales que $(\lambda_0, \lambda) \neq (0, 0)$ y satisfacen*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.

ii. *Si* $G(x) := \lambda_0 f(x) + \langle \lambda, g(x) \rangle$ *entonces* $G'(x_0) = 0$.

Claramente, si en este teorema x_0 es un punto normal con respecto a S, entonces $\lambda_0 > 0$ y los multiplicadores pueden ser elegidos de forma que $\lambda_0 = 1$. Por tanto, el Teorema 1.2.10 permitiría obtener el resultado del Teorema 1.2.3 reemplazando la palabra "regular" por "normal". Notemos también que esta prueba no requiere del Lema 1.2.2, el cual fue fundamental en la prueba de los multiplicadores de Lagrange de primer orden basada en la noción de regularidad.

Nota 1.2.11. *Observemos que, de acuerdo con esta última definición de normalidad, x_0 es un punto normal con respecto a $S_1(\lambda)$ si $\mu = 0$ es la única solución de*

i. $\mu_\alpha \geq 0$ *y* $\mu_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A,\ \lambda_\alpha = 0)$.

ii. $\sum_1^m \mu_i g_i'(x_0) = 0$.

Al igual que en el caso de normalidad con respecto a S, se puede probar que la normalidad con respecto a $S_1(\lambda)$ es equivalente a la siguiente condición:

(SMF) *Mangasarian-Fromovitz estricto en x_0.*

Si $\Gamma = \{\alpha \in I(x_0) \mid \lambda_\alpha > 0\}$, el conjunto $\{g_\beta'(x_0) \mid \beta \in B \cup \Gamma\}$ es linealmente independiente y $\exists h \in \mathbb{R}^n$ tal que

$$g_\alpha'(x_0; h) < 0 \ (\alpha \in I(x_0),\ \lambda_\alpha = 0), \quad g_\beta'(x_0; h) = 0 \ (\beta \in B \cup \Gamma).$$

Algunos autores llaman a la condición de arriba "calificación de restricción estricta de Mangasarian-Fromovitz", (ver [26]). Sin embargo (ver [44]), otros se mantienen renuentes a utilizar esta terminología porque, como ya mencionamos, el conjunto S_1 depende del multiplicador λ, el cual a su vez está determinado por la función f y no sólo por las restricciones. En adelante, nos referiremos a ella como la condición estricta de Mangasarian-Fromovitz.

Combinando los Teoremas 1.2.4 y 1.2.9, así como la Nota 1.2.11, podemos probar el siguiente resultado, *básico* en la teoría de condiciones necesarias de segundo orden.

Teorema 1.2.12. *Supongamos que existe $\lambda \in \mathbb{R}^m$ tal que*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.

ii. *Si* $F(x) := f(x) + \langle \lambda, g(x) \rangle$ *entonces* $F'(x_0) = 0$.

Si x_0 resuelve N(S) localmente y es un punto normal con respecto a $S_1(\lambda)$, entonces $F''(x_0, h) \geq 0$ para todo $h \in R_{S_1}(x_0)$.

La prueba de la equivalencia entre normalidad con respecto a S (S_1) y la condición (estricta) de Mangasarian-Fromovitz que se puede encontrar en [25] y que tiene como consecuencia los Teoremas 1.2.9 y 1.2.12 es extensa y complicada pues emplea propiedades fundamentales del cono tangente de S en x_0, mientras que Kyparisis proporciona en [26] una prueba más simple y directa. Dicha demostración se basa en el Teorema de alternativas de Motzkin (ver [23, Teorema 19, pág. 68]) el cual, a su vez, es consecuencia del lema de Farkas-Minkowski. Este enfoque, así como otra implicación particular de **(SMF)**, serán discutidos en la siguiente sección.

Pero, realmente ¿cuál es la relación entre estas nociones de normalidad?, pues hemos visto que ambas implican la regularidad de x_0 con respecto a S_1. Dado $x_0 \in S$, consideremos el conjunto de restricciones de igualdad

$$S_0 := \{x \in \mathbb{R}^n \mid g_i(x) = 0 \ (i \in I(x_0) \cup B)\}$$

y el cono linealizado de S_0 en x_0

$$R_{S_0}(x_0) = \{h \in \mathbb{R}^n \mid g_i'(x_0, h) = 0 \ (i \in I(x_0) \cup B)\}.$$

Como se verifica fácilmente, si x_0 resuelve nuestro problema original localmente, entonces resuelve también aquel determinado sólo por restricciones de igualdad para índices activos.

Nota 1.2.13. *Si x_0 es una solución local de $N(S)$ entonces es una solución local de $N(S_0)$.*

Demostración. Si $g_\alpha(x_0) < 0$, sea $\epsilon_\alpha > 0$ tal que $|x - x_0| < \epsilon_\alpha \Rightarrow g_\alpha(x) < 0$ y sea

$$N_0 := \{x \in \mathbb{R}^n : |x - x_0| < \epsilon\} \quad \text{donde } \epsilon = \text{mín}\{\epsilon_\alpha \mid g_\alpha(x_0) < 0\}.$$

Si $A = I(x_0)$, tomamos $N_0 := \mathbb{R}^n$. Como $S_0 \cap N_0 \subset S$, x_0 también minimiza localmente a f en S_0. □

Aplicando la definición de normalidad 1.2.8 a S_0, obtenemos lo siguiente.

Definición 1.2.14. *Un punto x_0 es normal con respecto a S_0 si $\lambda = 0$ es la única solución del sistema*

i. $\lambda_\alpha g_\alpha(x_0) = 0 \ (\alpha \in A)$.
ii. $\sum_1^m \lambda_i g_i'(x_0) = 0$.

Esto es equivalente a la siguiente condición.

(LI) *Condición de independencia lineal en x_0.*
El conjunto $\{g_i'(x_0) \mid i \in B \cup I(x_0)\}$ es linealmente independiente.

La cual es precisamente la definición de s-normalidad con respecto a S. Así, usando la Definición 1.2.14 es fácil ver que, si x_0 es un punto normal con respecto a S_0 (es decir, s-normal con respecto a S), entonces x_0 es un punto normal con respecto a S_1, lo que, a su vez, implica que es normal con respecto a S. Por tanto, la definición de normalidad 1.2.8 nos permite debilitar la hipótesis de s-normalidad con respecto a S y asegurar la regularidad con respecto a S_1, es decir, el Teorema 1.2.12 es un resultado más fuerte que el Teorema 1.2.7.

Finalizamos esta sección con algunos ejemplos que ilustran algunas de las características más importantes de los resultados obtenidos hasta ahora. En el siguiente ejemplo se muestra que, para una solución x_0 de $N(S)$, es posible tener $F''(x_0, h) < 0$ para alguna $h \in R_S(x_0)$ aun si la solución es normal con respecto a S_0 (y en consecuencia con respecto a S_1).

Ejemplo 1.2.1. *Considere el problema de minimizar $f(x_1, x_2) = x_1$ sujeto a*

$$g_1(x_1, x_2) = -x_1^2 - x_2 \leq 0, \quad g_2(x_1, x_2) = x_1 + x_2 = 0.$$

Claramente, $x_0 = (1, -1)$ es una solución local normal con respecto a S_0, pues $I(x_0) = \{1, 2\}$ y los gradientes $g_1'(x_0) = (-2, -1)$ y $g_2'(x_0) = (1, 1)$ son linealmente independientes. Luego

$$F(x_1, x_2) = x_1 + \lambda_1(-x_1^2 - x_2) + \lambda_2(x_1 + x_2)$$

y por tanto

$$F'(x_1, x_2) = (1 - 2\lambda_1 x_1 + \lambda_2, -\lambda_1 + \lambda_2),$$

de modo que $F'(x_0) = 0$ implica $\lambda_1 = \lambda_2 = 1$. Además,

$$F''(x_1, x_2) = \begin{pmatrix} -2\lambda_1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$$

de donde $F''(x_0; h, k) = -2h^2$. Observemos que

$$R_{S_0}(x_0) = R_{S_1}(x_0) = \{(h, k) \in \mathbb{R}^2 \mid -2h - k = 0, \ h + k = 0\} = \{(0, 0)\}$$

mientras que

$$R_S(x_0) = \{(h, k) \in \mathbb{R}^2 \mid -2h - k \leq 0, \ h + k = 0\}.$$

Por lo tanto, $(1, -1) \in R_S(x_0)$ y $F''(x_0; 1, -1) = -2 < 0$. $\square$

Veamos ahora un caso más sencillo donde la solución es global.

Ejemplo 1.2.2. *Considere el problema de minimizar $f(x) = -x^3$ sujeto a $g(x) = x - 1 \leq 0$.*

En este caso tenemos que $x_0 = 1$ es una solución global, normal con respecto a S_0, pues $g'(x_0) = 1 \neq 0$. Por otro lado,

$$F(x) = -x^3 + \lambda x - \lambda$$

y, por tanto, $F'(x) = -3x^2 + \lambda$ por lo que, si $F'(x_0) = 0$, se debe tener $\lambda = 3$. Como

$$R_S(x_0) = \{h \in \mathbb{R} \mid g'(x_0; h) = h \leq 0\},$$

se tiene que $-1 \in R_S(x_0)$ y $F''(x_0, -1) = -6 < 0$. $\square$

1.3. Normalidad y unicidad de los multiplicadores de Lagrange

En esta sección hacemos un breve análisis de la relación entre la normalidad de una solución de $N(S)$ y la unicidad de los multiplicadores asociados a ésta.

En [23, Teorema 3.10.4], se establece un resultado sobre la unicidad de los multiplicadores de Lagrange asociados a una solución del problema $N(S)$ como parte de un teorema donde se derivan condiciones de segundo orden. Para explicar la relación entre estos dos aspectos (unicidad de multiplicadores y condiciones de segundo orden) debemos mencionar que, de acuerdo con [23], en [4] y [10] se establece un resultado falso sobre condiciones necesarias de segundo orden. Ahí, los autores afirman que, suponiendo **(LI)**, las condiciones de segundo orden se cumplen en todo el conjunto $R_S(x_0)$, afirmación que contradicen los ejemplos 1.2.1 y 1.2.2.

En [26], siguiendo la "prueba" de este resultado falso que da Ben-Tal en [10, Teorema 3.3], Kyparisis establece un resultado correcto, imponiendo una condición a la que llama "calificación de restricción estricta de Mangasarian-Fromovitz" (introducida por Fujiwara, Han y Mangasarian en [21]) y prueba que las condiciones de segundo orden se satisfacen en el subcono de $R_S(x_0)$ que considera el signo de los multiplicadores de Lagrange $\{\lambda_i\}_1^m$ correspondientes (o sea, $R_{S_1}(x_0)$). Más aún, en el mismo artículo, Kyparisis muestra que la condición **(SMF)** es también equivalente a la unicidad del multiplicador λ que satisface las condiciones del Teorema 1.2.3 y que define a S_1.

Definición 1.3.1. *Denotamos por* $\Lambda(f, x_0)$ *al conjunto de todos los* $\lambda \in \mathbb{R}^n$ *que satisfacen la condición de Kuhn-Tucker o regla de los multiplicadores de Lagrange de primer orden (Teorema 1.2.3), o sea,*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.

ii. *Si* $F(x) := f(x) + \langle \lambda, g(x) \rangle$ *entonces* $F'(x_0) = 0$.

En [26] se mencionan, en particular, las calificaciones de restricción **(MF)** y **(LI)**. Hemos dicho que estas condiciones son equivalentes a la normalidad con respecto a S y con respecto a S_0 respectivamente. Recordemos también que, dado $\lambda \in \Lambda(f, x_0)$, la normalidad con respecto a $S_1(\lambda)$ es equivalente a la condición **(SMF)**, la cual es más restrictiva que **(MF)**, pero menos que **(LI)**. Para esta última se tiene el siguiente resultado.

Proposición 1.3.2. *Si* x_0 *es una solución local de* $N(S)$ *que satisface la condición* **(LI)**, *entonces existe un único* $\lambda \in \Lambda(f, x_0)$.

Demostración. La existencia es consecuencia de los Teoremas 1.2.3 y 1.2.6. Luego, si $\mu, \lambda \in \Lambda(f, x_0)$, entonces

i. $(\lambda - \mu)_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.

ii. $\sum_1^m (\lambda - \mu)_i g_i'(x_0) = 0$

y, por la independencia lineal de los $g_i'(x_0)$ $(i \in I(x_0) \cup B)$, $\mu = \lambda$. □

Por otro lado, la caracterización de la unicidad de multiplicadores de Lagrange dada en [26] corresponde al siguiente resultado.

Teorema 1.3.3. *Sea* $x_0 \in S$ *para el cual existe* $\lambda \in \Lambda(f, x_0)$. *Entonces* **(SMF)** *se cumple en* $x_0 \Leftrightarrow \Lambda(f, x_0) = \{\lambda\}$.

Un segundo resultado de [26], que relaciona **(SMF)** con condiciones de segundo orden, establece lo siguiente.

Teorema 1.3.4. *Supongamos que* x_0 *resuelve N(S) localmente y* $\lambda \in \Lambda(f, x_0)$. *Si* $\Lambda(f, x_0) = \{\lambda\}$, *entonces* $F''(x_0, h) \geq 0$ *para todo* h *que satisface*

$$g_\alpha'(x_0, h) < 0 \ (\alpha \in I(x_0), \lambda_\alpha = 0), \quad g_\beta'(x_0, h) = 0 \ (\beta \in B \cup \Gamma).$$

Considerando lo anterior, el Teorema 1.2.12, el resultado principal de este capítulo, puede ser reformulado de la siguiente manera.

Teorema 1.3.5. *Supongamos que x_0 resuelve N(S) localmente y $\Lambda(f, x_0) = \{\lambda\}$. Entonces $F''(x_0, h) \geq 0$ para todo $h \in R_{S_1}(x_0)$.*

El Teorema 1.3.3 dio lugar también a una controversia sobre el concepto de "calificación de restricción". En [44], Wachsmuth llama así a las hipótesis impuestas a las restricciones que garantizan que la condición $-f'(x_0) \in R_S^*(x_0)$ sea una condición necesaria para optimalidad en nuestro problema.

Como hemos mencionado, en [23] usando el lema de Farkas y en [24, 25] la teoría de conos convexos, se muestra que

$$\Lambda(f, x_0) \neq \emptyset \Leftrightarrow f'(x_0, h) \geq 0 \text{ para toda } h \in R_S(x_0)$$

(es decir, $-f'(x_0) \in R_S^*(x_0)$). Ahora bien, por el Teorema 1.0.2 sabemos que, si x_0 es solución local, entonces

$$f'(x_0, h) \geq 0 \text{ para toda } h \in T_S(x_0)$$

(es decir, $-f'(x_0) \in T_S^*(x_0)$) y, como $R_S^*(x_0) \subset T_S^*(x_0)$, vemos que la definición de Wachsmuth coincide con nuestra definición previa de calificación de restricciones.

Como se destaca en [44], las calificaciones de restricción son independientes de la función objetivo f. Por lo tanto, si una calificación de restricción implica cierta propiedad de los elementos $\lambda \in \Lambda(f, x_0)$, dicha propiedad debería satisfacerse para todas las funciones objetivo (que alcanzan un mínimo local en x_0). Considerando lo anterior, Wachsmuth define en [44]

$$\mathcal{F}(x_0) := \{f \in C^1(\mathbb{R}^n, \mathbb{R}) \mid x_0 \text{ minimiza a } f \text{ localmente en } S\}.$$

El resultado sobre unicidad de los multiplicadores de Lagrange dado en [44] es el siguiente.

Teorema 1.3.6. *Sea $x_0 \in S$. Entonces $\Lambda(f, x_0)$ es unitario para toda $f \in \mathcal{F}(x_0) \Leftrightarrow$* **(LI)** *se satisface en x_0.*

La diferencia principal entre este resultado y el Teorema 1.3.3 es, como mencionamos en la sección anterior, que la condición **(SMF)** presupone la existencia de multiplicadores de Lagrange y depende (indirectamente) de la función objetivo f, por lo cual Wachsmuth no la considera una calificación de restricción y hace notar, en cambio, que **(LI)** es una calificación de restricción que asegura la existencia y unicidad de multiplicadores de Lagrange para toda $f \in \mathcal{F}(x_0)$.

Nuestro último ejemplo muestra una solución x_0 regular con respecto a S y $\lambda \in \Lambda(f, x_0)$ para los cuales existe $h \in R_{S_1(\lambda)}(x_0)$ tal que $F''(x_0, h) < 0$.

Ejemplo 1.3.1. *Considere el problema de minimizar $f(x,y) = y + x^2$ en el conjunto*

$$S = \{(x,y) \mid g_\alpha(x,y) \leq 0 \ (\alpha = 1,2,3)\}$$

donde $g_1(x,y) = -y$, $g_2(x,y) = \cos x - y$, $g_3(x,y) = \cos 2x - y$.

Es fácil verificar que $(0,1)$ resuelve el problema. El conjunto de índices activos en este punto es $I(0,1) = \{2,3\}$, y es un punto regular con respecto S pues

$$T_S(0,1) = \{(h,k) \mid k \geq 0\}, \quad R_S(0,1) = \{(h,k) \mid -k \leq 0\}.$$

Por el Teorema 1.2.3, existe $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ con $\lambda_1 = 0$ y $\lambda_2, \lambda_3 \geq 0$ tal que, si

$$\begin{aligned} F(x,y) &= f(x,y) + \langle \lambda, g(x,y) \rangle \\ &= y + x^2 - (\lambda_1 + \lambda_2 + \lambda_3)y + \lambda_2 \cos x + \lambda_3 \cos 2x \end{aligned}$$

entonces $F'(0,1) = 0$. Como

$$F'(x,y) = (2x - \lambda_2 \operatorname{sen} x - 2\lambda_3 \operatorname{sen} 2x, 1 - \lambda_1 - \lambda_2 - \lambda_3)$$

tenemos que, para $\lambda_1 = 0$, $F'(0,1) = (0, 1 - \lambda_2 - \lambda_3)$, de donde $\lambda_2 + \lambda_3 = 1$. Esta última ecuación tiene múltiples soluciones $\lambda_2, \lambda_3 \geq 0$, por lo cual $(0,1)$ no satisface la condición de normalidad con respecto a $S_1(\lambda)$, para ninguna de ellas.

Notemos ahora que

$$F''(x,y) = \begin{pmatrix} 2 - \lambda_2 \cos x - 4\lambda_3 \cos 2x & 0 \\ 0 & 0 \end{pmatrix}$$

de manera que

$$F''(0,1;h,k) = (2 - \lambda_2 - 4\lambda_3)h^2.$$

Así, eligiendo $\lambda_2 = 0$ y $\lambda_3 = 1$, obtenemos que

$$R_{S_1}(0,1) = \{(h,k) \mid -k \leq 0, \ k = 0\}$$

y por tanto, para cualquier $h \neq 0$, $(h, 0) \in R_{S_1}(0, 1)$ pero $F''(0, 1; h, 0) = -2h^2 < 0$.

Capítulo 2

Problemas de optimización con restricciones isoperimétricas

En este capítulo abordaremos el problema de Lagrange con puntos extremos fijos en cálculo de variaciones y en control óptimo, definidos sobre arcos suaves a trozos y controles continuos a trozos sujetos a restricciones en forma de igualdades y desigualdades. En la literatura clásica donde se estudian estos problemas (ver por ejemplo, [3, 22, 24, 25]), podemos encontrar condiciones necesarias de segundo orden que son válidas en un conjunto de direcciones críticas que consideran igualdades y desigualdades dependiendo del signo de los multiplicadores de Lagrange del extremo en consideración, pero que requieren de una hipótesis de normalidad con respecto a un conjunto definido sólo por restricciones de igualdad para índices activos, condición que además asegura la unicidad del multiplicador de Lagrange correspondiente.

Recientemente, en [9], se probó que en el contexto de cálculo de variaciones es posible garantizar que se satisfacen condiciones de segundo orden en el mismo conjunto de direcciones bajo una hipótesis considerablemente más débil, estableciendo así un resultado para este problema análogo al del Teorema 1.2.12, mientras que en [17] mostramos que esta nueva condición es suficiente también para que el multiplicador asociado a una solución sea único y establecimos resultados análogos a los Teoremas 1.3.2 y 1.3.6.

En la primera sección de este apartado hacemos una breve exposición de los resultados mencionados arriba. Para el problema en el contexto de control óptimo (que abordamos en la segunda sección) ocurre una situación similar. En [5] se estableció que la contraparte del Teorema 1.2.12 es también válida. Sin embargo, no se responde a la pregunta si la nueva hipótesis garantiza la unicidad del multiplicador, pregunta que respondemos mostrando que no sólo es suficiente, sino equivalente.

2.1. Cálculo de variaciones

Para establecer el problema, supongamos que tenemos dados un intervalo $T := [t_0, t_1]$ en $\mathbb{R}$, dos puntos fijos ξ_0, ξ_1 en $\mathbb{R}^n$, funciones L y L_γ ($\gamma = 1, \ldots, q$) de $T \times \mathbb{R}^n \times \mathbb{R}^n$ a $\mathbb{R}$. Denotemos por X al espacio de funciones C^1 a trozos que van de T a $\mathbb{R}^n$ y sean

$$X_e := \{x \in X \mid x(t_0) = \xi_0,\ x(t_1) = \xi_1\},$$

$$S := \{x \in X_e \mid I_\alpha(x) \leq 0\ (\alpha \in R),\ I_\beta(x) = 0\ (\beta \in Q)\},$$

donde $R = \{1, \ldots, r\}$, $Q = \{r + 1, \ldots, q\}$,

$$I_\gamma(x) = \int_{t_0}^{t_1} L_\gamma(t, x(t), \dot{x}(t))dt \quad (x \in X).$$

El problema, que denotaremos como (V), consiste en minimizar I sobre S, donde

$$I(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \quad (x \in X).$$

Los elementos de X se conocen como *arcos*, los de S como *arcos admisibles* y diremos que un arco admisible x es una solución local del problema (V) si (en caso necesario, cambiando $T \times \mathbb{R}^n$ por $T \times U$, donde U es un abierto de $\mathbb{R}^n$) cualquier otro arco admisible y satisface que $I(x) \leq I(y)$.

Para $x \in X$, denotaremos $(t, x(t), \dot{x}(t))$ por $(\tilde{x}(t))$, y supondremos que las funciones L, L_γ son C^1 o C^2 si aparecen segundas derivadas.

Para todo $x \in X$, se define la *primera variación de I en x en la dirección de y* como

$$I'(x; y) := \int_{t_0}^{t_1} \{L_x(\tilde{x}(t))y(t) + L_{\dot{x}}(\tilde{x}(t))\dot{y}(t)\}dt \quad (y \in X)$$

y la *segunda variación de I en x en la dirección de y* como

$$I''(x; y) := \int_{t_0}^{t_1} 2\Omega(t, y(t), \dot{y}(t))dt \quad (y \in X)$$

donde, para todo $(t, y, \dot{y}) \in T \times \mathbb{R}^n \times \mathbb{R}^n$,

$$2\Omega(t, y, \dot{y}) := \langle y, L_{xx}(\tilde{x}(t))y\rangle + 2\langle y, L_{x\dot{x}}(\tilde{x}(t))\dot{y}\rangle + \langle \dot{y}, L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}\rangle.$$

En las definiciones anteriores L_x y $L_{\dot{x}}$ representan, respectivamente, a la primer derivada parcial de L con respecto a x y con respecto a $\dot{x}$, no a los integrandos de las integrales I_γ los cuales denotamos con subíndices griegos. Asimismo, L_{xx}, $L_{x\dot{x}}$ y $L_{\dot{x}\dot{x}}$ representan las segundas derivadas parciales.

La primera y segunda variación de otras integrales como I_γ se definen de manera similar. Definimos el conjunto de *variaciones admisibles* como

$$Y := \{y \in X \mid y(t_0) = y(t_1) = 0\}.$$

El siguiente resultado puede considerarse el análogo del Teorema 1.2.10, la condición de Fritz John, pues provee condiciones de primer orden para (V) (ver [24]).

Teorema 2.1.1. *Supongamos que $x_0 \in S$ resuelve (V). Entonces existen $\lambda_0 \geq 0$ y $\lambda \in \mathbb{R}^q$ tales que $(\lambda_0, \lambda) \neq (0,0)$ y*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R)$.

ii. *Si* $J_0(x) := \lambda_0 I(x) + \sum_1^q \lambda_\gamma I_\gamma(x)$, *entonces* $J_0'(x_0; y) = 0$ *para toda* $y \in Y$.

Basados en este resultado, definimos la normalidad con respecto a S (compare con 1.2.8) o *normalidad débil*, como sigue.

Definición 2.1.2. *Un arco $x_0 \in S$ se llamará normal con respecto a S o débilmente normal si $\lambda = 0$ es la única solución de*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R)$.

ii. $\sum_1^q \lambda_\gamma I_\gamma'(x_0; y) = 0$ *para toda* $y \in Y$.

Observemos que, si x_0 resuelve (V) y es débilmente normal, entonces $\lambda_0 > 0$ en el Teorema 2.1.1. Así, los multiplicadores pueden elegirse de forma que $\lambda_0 = 1$. En este caso, la pareja $(x_0, \lambda) \in S \times \mathbb{R}^q$ se llamará *extremo* y al conjunto de extremos lo denotamos por $\mathcal{E}$.

Definición 2.1.3. *Denotemos por $\Lambda(L, x_0)$ al conjunto de los multiplicadores de Lagrange asociados a x_0, es decir, a todos los $\lambda \in \mathbb{R}^q$ tales que $(x_0, \lambda) \in \mathcal{E}$, i.e.,*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R)$.

ii. *Si* $J(x) := I(x) + \sum_1^q \lambda_\gamma I_\gamma(x)$, *entonces* $J'(x_0; y) = 0$ *para toda* $y \in Y$.

La condición (ii) es equivalente a la existencia de $c \in \mathbb{R}^n$ tal que

$$F_{\dot{x}}(\tilde{x}(t)) = \int_{t_0}^{t} F_x(\tilde{x}(s))ds + c \quad (t \in T)$$

donde $F := L + \sum_1^q \lambda_\gamma L_\gamma$ (ver [14, 24]).

Como ocurre en el caso de programación no lineal, la normalidad con respecto a S de una solución local implica la existencia de multiplicadores de Lagrange, mas no la unicidad de éstos. Para ello se requiere de una hipótesis más fuerte, al igual que para establecer condiciones necesarias de segundo orden para (V).

Para introducir esta hipótesis, consideremos el conjunto de índices activos en un arco admisible x_0, el cual denotamos como

$$A(x_0) = \{\alpha \in R \mid I_\alpha(x_0) = 0\}$$

y, como en el capítulo previo, definamos al conjunto

$$S_0 := S_0(x_0) = \{x \in X_e \mid I_\gamma(x_0) = 0 \ (\gamma \in A(x_0) \cup Q)\}.$$

Notemos que, por definición, x_0 es normal con respecto a S_0 si $\lambda = 0$ es la única solución de

i. $\lambda_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R)$.
ii. $\sum_1^q \lambda_\gamma I'_\gamma(x_0; y) = 0$ para toda $y \in Y$.

Claramente, esta condición, que llamaremos *normalidad fuerte*, es equivalente a la independencia lineal de las primeras variaciones de $I_\gamma(x_0)$ $(\gamma \in A(x_0) \cup Q)$ en el conjunto de direcciones Y, lo que a su vez es equivalente a la existencia de $y_\gamma \in Y$ $(\gamma \in A(x_0) \cup Q)$ tales que el determinante

$$|I'_\beta(x_0; y_\gamma)| \quad (\beta, \gamma \in A(x_0) \cup Q)$$

es diferente de cero.

En textos clásicos, como [14, 24], las condiciones de primer y segundo orden para (V) se establecen de la siguiente manera.

Teorema 2.1.4. *Si x_0 resuelve (V) y es fuertemente normal, entonces existe un único $\lambda \in \mathbb{R}^q$ tal que $\lambda \in \Lambda(L, x_0)$.*

Daremos una prueba del resultado anterior al final de esta sección.

Teorema 2.1.5. *Supongamos que $\lambda \in \Lambda(L, x_0)$. Si x_0 resuelve (V) y es fuertemente normal, entonces $J''(x_0; y) \geq 0$ para toda $y \in Y$ que satisfaga*

a. $I'_\alpha(x_0; y) \leq 0$ $(\alpha \in A(x_0),\ \lambda_\alpha = 0)$;

b. $I'_\beta(x_0; y) = 0$ *($\beta \in R$ con $\lambda_\beta > 0$ o $\beta \in Q$).*

Observemos que en los dos resultados se requiere de la condición de normalidad fuerte. Sin embargo, en un artículo reciente [9, Teorema 1.5], se obtuvo la misma condición de segundo orden del Teorema 2.1.5 pero bajo una hipótesis más débil. El resultado obtenido es análogo al Teorema 1.2.12, por lo que la hipótesis se expresa en términos del conjunto

$$S_1(\lambda) := \{x \in S \mid J(x) = I(x)\},$$

para $\lambda \in \Lambda(L, x_0)$ dado y J definida por $J(x) := I(x) + \sum_1^q \lambda_\gamma I_\gamma(x)$. Explícitamente,

$$\begin{aligned} S_1\ (= S_1(\lambda)) = \{x \in X_e \mid I_\alpha(x) \leq 0\ (\alpha \in R,\ \lambda_\alpha = 0), \\ I_\beta(x) = 0\ (\beta \in R \text{ con } \lambda_\beta > 0,\ \text{ o } \beta \in Q)\}. \end{aligned}$$

Vemos que x_0 es normal con respecto a $S_1(\lambda)$ si $\mu = 0$ es la única solución para

i. $\mu_\alpha \geq 0$ y $\mu_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R,\ \lambda_\alpha = 0)$.

ii. $\sum_1^q \mu_\gamma I'_\gamma(x_0; y) = 0$ para toda $y \in Y$.

Denotando por $R_{S_1}(x_0)$ a los elementos de Y que satisfacen las condiciones (a) y (b) del Teorema 2.1.5, es decir, las restricciones tangenciales de S_1 en x_0, el Teorema 1.5 de [9] establece lo siguiente.

Teorema 2.1.6. *Supongamos que x_0 resuelve (V) y existe $\lambda \in \Lambda(L, x_0)$. Si x_0 es normal con respecto a $S_1(\lambda)$, entonces $J''(x_0; y) \geq 0$ para toda $y \in R_{S_1}(x_0)$.*

La prueba de este resultado es análoga a la del Teorema 1.2.12 que proporciona Hestenes en [25] y está basada en dos resultados fundamentales relacionados con la condición de normalidad. El primero [9, Proposición 2.3] es una caracterización en términos de una condición que los autores llaman "propiedad" y que enunciamos a continuación.

Definición 2.1.7. *Se dice que $x_0 \in S$ es propio con respecto a S si*

a. $\{I'_\beta(x_0; \cdot) \mid \beta \in Q\}$ *es linealmente independiente en Y.*

b. $\exists y \in Y$ *tal que* $I'_\alpha(x_0; y) < 0$ $(\alpha \in A(x_0))$ *y* $I'_\beta(x_0; y) = 0$ $(\beta \in Q)$.

Es claro que la condición de "propiedad" corresponde a una condición de tipo Mangasarian-Fromovitz.

Sustituyendo S por S_1, la caracterización anterior establece el siguiente resultado.

Proposición 2.1.8. *Sea $x_0 \in S$ para el cual existe $\lambda \in \Lambda(L, x_0)$, y sea $\Gamma := \{\alpha \in A(x_0) \mid \lambda_\alpha > 0\}$. Entonces x_0 es normal con respecto a $S_1(\lambda)$ si y sólo si es propio con respecto a $S_1(\lambda)$, es decir, si*

a. $\{I'_\beta(x_0; \cdot) \mid \beta \in \Gamma \cup Q\}$ *es linealmente independiente en Y.*

b. $\exists y \in Y$ *tal que* $I'_\alpha(x_0; y) < 0$ $(\alpha \in A(x_0) - \Gamma)$ *y* $I'_\beta(x_0; y) = 0$ $(\beta \in \Gamma \cup Q)$.

Para enunciar el otro resultado básico establecido en [9], que conduce al Teorema 2.1.6, necesitaremos la siguiente definición.

Definición 2.1.9. *Denotamos por $T_S(x_0)$ al cono de tangentes secuenciales, determinado por los arcos unitarios $y \in Y$ para los cuales existe una sucesión $\{x_q\}$ en S que converge a x_0 en la dirección y, es decir,*

$$\lim_{k\to\infty} \|x_k - x_0\| = 0, \quad \lim_{k\to\infty} \frac{x_k - x_0}{\|x_k - x_0\|} = y. \tag{2.1.1}$$

La norma en la Definición 2.1.9 es la norma débil en X, dada por

$$\|x\| := \sup\{|x(t)|^2 + |\dot{x}(t)|^2\}^{1/2},$$

pues considerando esta norma, si $\{x_k\}$ converge a x_0 en la dirección y, entonces

$$\lim_{k\to\infty} \frac{I(x_k) - I(x_0)}{\|x_k - x_0\|} = I'(x_0; y). \tag{2.1.2}$$

$$\lim_{k\to\infty} \frac{I(x_k) - I(x_0) - I'(x_0; x_k - x_0)}{\|x_k - x_0\|^2} = \frac{1}{2} I''(x_0; y). \tag{2.1.3}$$

Note que la ecuación 2.1.2 implica que $T_S(x_0) \subset R_S(x_0)$. Diremos que x_0 es un *arco regular con respecto a S* si $T_S(x_0) = R_S(x_0)$.

Ahora, supongamos que x_0 es una solución local de (V) tal que $\lambda \in \Lambda(L, x_0)$ y sea $y \in T_{S_1}(x_0)$, sin pérdida de generalidad, podemos suponer que y es un arco unitario, por lo tanto, existe una sucesión $\{x_k\} \subset S_1$ que converge a x_0 en la dirección y. Luego, por la definición de $\Lambda(L, x_0)$, sabemos que $J'(x_0; y) = 0$ y que $J(x) = I(x)$ para $x \in S_1$, por lo tanto, para k suficientemente grande, se debe cumplir que $J(x_k) \geq J(x_0)$, implicándose por la ecuación 2.1.3 que $J''(x_0; y) \geq 0$. La conclusión del Teorema 2.1.6 se sigue entonces sustituyendo S por S_1 en [9, Teorema 2.4] que establece lo siguiente.

Teorema 2.1.10. *Si x_0 es un arco normal con respecto a S, entonces es regular con respecto a S.*

Probaremos ahora que los teoremas sobre unicidad establecidos en [26] y [44] son válidos también para nuestro problema con restricciones isoperimétricas (V). Observe que, para este fin, basta que las funciones involucradas sean C^1.

Teorema 2.1.11. *Sea $x_0 \in S$ y supongamos que existe $\lambda \in \Lambda(L, x_0)$. Entonces las siguientes afirmaciones son equivalentes:*

a. *x_0 es normal con respecto a $S_1(\lambda)$.*

b. *λ es el único elemento de $\Lambda(L, x_0)$.*

Demostración.

(a) $\Rightarrow$ (b): Sea $\bar{\lambda} \in \Lambda(L, x_0)$ y definamos $\mu := \bar{\lambda} - \lambda$. Tenemos entonces que

$$\mu_\alpha = \bar{\lambda}_\alpha \geq 0 \quad \text{y} \quad \mu_\alpha I_\alpha(x_0) = \bar{\lambda}_\alpha I_\alpha(x_0) = 0 \ (\alpha \in R,\ \lambda_\alpha = 0)$$

y, si $\bar{J}(x) := I(x) + \sum_1^q \bar{\lambda}_\gamma I_\gamma(x)$, entonces

$$0 = \bar{J}'(x_0; y) - J'(x_0; y) = \sum_1^q \mu_\gamma I'_\gamma(x_0; y)$$

para toda $y \in Y$. Por (a), $\mu = 0$ y consecuentemente $\bar{\lambda} = \lambda$.

(b) $\Rightarrow$ (a): $\neg$(a) $\Rightarrow$ $\neg$(b): Supongamos que x_0 no es un arco normal con respecto a $S_1(\lambda)$. Entonces, podemos encontrar $\mu \in \mathbb{R}^q$, $\mu \neq 0$ satisfaciendo

i. $\mu_\alpha \geq 0$ y $\mu_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R,\ \lambda_\alpha = 0)$;

ii. $\sum_1^q \mu_\gamma I'_\gamma(x_0; y) = 0$ para toda $y \in Y$.

Notemos que, si $\Gamma \neq \emptyset$, podemos tomar a μ de forma que

$$\text{máx}\{|\mu_\alpha| : \alpha \in \Gamma\} < \text{mín}\{\lambda_\alpha : \alpha \in \Gamma\} \tag{2.1.4}$$

pues, si para todo $\alpha \in \Gamma$ se tiene que $\mu_\alpha = 0$, la desigualdad se cumple de forma trivial. De lo contrario, $0 < M := \text{máx}\{|\mu_\alpha| : \alpha \in \Gamma\}$ y, si $m := \text{mín}\{\lambda_\alpha : \alpha \in \Gamma\}$, para todo $k \in (0, m)$, $\bar{\mu} := k\mu/M$, es una solución no trivial de i. y ii. que, claramente, satisface la desigualdad (2.1.4).

Definimos $\hat{\lambda} := \lambda + \mu$. Mostraremos que $\hat{\lambda} \in \Lambda(L, x_0)$, implicando $\neg$(b) pues

$\hat\lambda \neq \lambda$. En efecto, si $\hat J(x) := I(x) + \sum_1^q \hat\lambda_\gamma I_\gamma(x)$, entonces

$$\hat J'(x_0;y) = J'(x_0;y) + \sum_1^q \mu_\gamma I'_\gamma(x_0;y) = 0 \quad (y \in Y)$$

y $\hat\lambda$ satisface 2.1.3(ii). Para ver 2.1.3(i), consideremos $\alpha \in R$. Si $I_\alpha(x_0) = 0$ entonces $\hat\lambda_\alpha I_\alpha(x_0) = 0$. Si $I_\alpha(x_0) < 0$, por 2.1.3(i) con respecto a λ concluimos que $\lambda_\alpha = 0$ y, por la propiedad (i) de μ, se sigue que $0 = \mu_\alpha I(x_0) = \hat\lambda_\alpha I(x_0)$. Finalmente, si $\lambda_\alpha = 0$, entonces $\hat\lambda_\alpha = \mu_\alpha \geq 0$. Si $\lambda_\alpha > 0$, entonces

$$\hat\lambda_\alpha = \lambda_\alpha + \mu_\alpha \geq \min_{i\in\Gamma} \lambda_i + \mu_\alpha > \max_{i\in\Gamma} |\mu_i| + \mu_\alpha \geq 0,$$

lo que prueba que $\hat\lambda \in \Lambda(L, x_0)$. □

Este resultado y el Teorema 2.1.6 conducen a la siguiente condición de segundo orden necesaria para optimalidad.

Corolario 2.1.12. *Sean $x_0 \in S$ y $\lambda \in \Lambda(L, x_0)$. Si x_0 resuelve (V) y $\Lambda(L, x_0) = \{\lambda\}$, entonces $J''(x_0; y) \geq 0$ para toda $y \in R_{S_1}(x_0)$.*

Para establecer la contraparte del Teorema 1.3.6, denotamos por $\mathcal{F}(x_0)$ al conjunto de todas las funciones L mapeando $T \times \mathbb{R}^n \times \mathbb{R}^n$ a $\mathbb{R}$ de clase C^1 tales que x_0 resuelve el problema V(L) de minimizar

$$I(x) = \int_{t_0}^{t_1} L(t, x(t), \dot x(t))dt$$

sobre S. El Teorema 1.3.6 [44, Teorema 2] corresponde, en el contexto de restricciones isoperimétricas, al siguiente resultado.

Teorema 2.1.13. *Para $x_0 \in S$ son equivalentes:*

a. *x_0 es fuertemente normal.*

b. *$\Lambda(L, x_0)$ es unitario para toda $L \in \mathcal{F}(x_0)$.*

Demostración.

(a) ⇒ (b): Notemos que este es el enunciado del Teorema 2.1.4. Para probarlo, considere $L \in \mathcal{F}(x_0)$. Como x_0 resuelve V(L), tenemos por (a) que $\Lambda(L, x_0) \neq \emptyset$ pues, en particular, x_0 es normal con respecto a S. Supongamos que λ y $\hat\lambda$ pertenecen a $\Lambda(L, x_0)$ y definamos $\mu := \lambda - \hat\lambda$. Por la Definición 2.1.3, μ satisface

i. $\mu_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R)$;

ii. $\sum_1^q \mu_\gamma I'_\gamma(x_0; y) = 0$ para toda $y \in Y$,
implicando, por (a), que $\mu = 0$.

(b) $\Rightarrow$ (a): Para todo $i \in R \cup Q$, definimos $\mu_i := 1$ si $i \in A(x_0)$ y $\mu_i := 0$ en otro caso, y

$$L(t, x, \dot{x}) := -\sum_1^q \mu_\gamma L_\gamma(t, x, \dot{x}).$$

Notemos que

$$I(x_0) = \int_{t_0}^{t_1} L(t, x_0(t), \dot{x}_0(t))dt = 0 \leq \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \quad (x \in S)$$

y, por tanto, $L \in \mathcal{F}(x_0)$. Para esta función, es claro que $(x_0, \mu) \in \mathcal{E}$, lo que implica que $\mu \in \Lambda(L, x_0)$. Luego, supongamos que $\nu \in \mathbb{R}^q$ satisface

i. $\nu_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R)$;

ii. $\sum_1^q \nu_\gamma I'_\gamma(x_0; y) = 0$ para toda $y \in Y$,
si mostramos que $\nu = 0$ habremos probado la condición (a). Para ello, definimos para todo $i \in R \cup Q$,

$$\hat{\mu}_i := \mu_i + \frac{\nu_i}{\beta} \quad \text{donde} \quad \beta = 1 + \text{máx}\{|\nu_\alpha| : \alpha \in A(x_0)\}.$$

Afirmamos que $(x_0, \hat{\mu}) \in \mathcal{E}$. En efecto, 2.1.3(i) se cumple pues, si $I_\alpha(x_0) = 0$, entonces

$$\hat{\mu}_\alpha = \frac{\beta + \nu_\alpha}{\beta} \geq 0$$

y, si $I_\alpha(x_0) < 0$, entonces $\hat{\mu}_\alpha I_\alpha(x_0) = 0$. Para 2.1.3(ii), observemos que si

$$J(x) = I(x) + \sum_1^q \hat{\mu}_\gamma I_\gamma(x),$$

como $(x_0, \mu) \in \mathcal{E}$, por la condición (ii) de ν se tiene que $J'(x_0; y) = 0$ para toda $y \in Y$, lo que prueba nuestra afirmación y por tanto, $\hat{\mu} \in \Lambda(L, x_0)$. Por (b) $\hat{\mu} = \mu$, implicando $\nu = 0$. $\square$

2.2. Control óptimo

En esta sección mostraremos que, pese a que la naturaleza del conjunto de multiplicadores de Lagrange asociados a una solución es diferente a los casos de cálculo de variaciones y dimensión finita, las caracterizaciones de la unicidad obtenidas por Kyparisis y Wachsmuth son igualmente válidas para el problema isoperimétrico de Lagrange con puntos extremos fijos en el contexto de control óptimo.

Dicho problema puede plantearse como sigue. Consideremos un intervalo $T := [t_0, t_1]$ en $\mathbb{R}$, dos puntos ξ_0, ξ_1 en $\mathbb{R}^n$, funciones L y L_γ ($\gamma = 1, \ldots, q$) que mapean $T \times \mathbb{R}^n \times \mathbb{R}^m$ a $\mathbb{R}$, f que mapea $T \times \mathbb{R}^n \times \mathbb{R}^m$ a $\mathbb{R}^n$, todas de clase C^1 en las variables (t, x, u).

Denotemos por X al espacio de funciones C^1 a trozos mapeando T a $\mathbb{R}^n$, que llamaremos *estados*, y por $\mathcal{U}$ al espacio de funciones continuas a trozos mapeando T a $\mathbb{R}^m$, denominadas *controles*. Definimos $Z := X \times \mathcal{U}$ y consideremos los siguientes conjuntos:

$$D := \{(x, u) \in Z \mid \dot{x}(t) = f(t, x(t), u(t)) \ \forall t \in T,\ x(t_0) = \xi_0,\ x(t_1) = \xi_1\},$$

$$S := \{(x, u) \in D \mid I_\alpha(x, u) \leq 0,\ I_\beta(x, u) = 0 \ (\alpha \in R,\ \beta \in Q)\},$$

donde $R = \{1, \ldots, r\}$, $Q = \{r + 1, \ldots, q\}$,

$$I_\gamma(x, u) = \int_{t_0}^{t_1} L_\gamma(t, x(t), u(t))dt \quad ((x, u) \in Z).$$

El problema, que denotaremos por (IP), consiste en minimizar el funcional I sobre S, donde

$$I(x, u) := \int_{t_0}^{t_1} L(t, x(t), u(t))dt \quad ((x, u) \in Z).$$

Los elementos de Z se conocen como *procesos*, conjunto en el que consideraremos la norma

$$\|z\| := \|(x, u)\| = \sup_{t \in T}(|x(t)| + |\dot{x}(t)|) + \sup_{t \in T} |u(t)|\,. \tag{2.2.1}$$

Llamaremos a los procesos de S *admisibles* y diremos que un proceso (x, u) es *solución de* o *resuelve* (IP) (localmente) si (x, u) es admisible e $I(x, u) \leq I(y, v)$ para todo (y, v) admisible en una vecindad de (x, u). Para simplificar la notación, $(\tilde{x}_0(t))$ representará a la terna $(t, x_0(t), u_0(t))$, sustituyendo $u_0(t)$ por el límite lateral que corresponda si t es un punto de discontinuidad de u_0.

El resultado clásico sobre condiciones necesarias de primer orden para (IP) es el Principio Máximo de Pontryagin, el cual se expresa en términos de la función hamiltoniano, definida como sigue:

$$H(t,x,u,p,\lambda,\lambda_0) := \langle p, f(t,x,u)\rangle - \lambda_0 L(t,x,u) - \sum_1^q \lambda_\gamma L_\gamma(t,x,u).$$

La versión del Principio Máximo que enunciamos a continuación corresponde a [24, pág. 254, Teorema 2.1].

Teorema 2.2.1. Principio Máximo de Pontryagin. *Si (x_0,u_0) resuelve (IP), entonces existen $\lambda_0 \geq 0$, $p \in X$ y $\lambda \in \mathbb{R}^q$ que no se anulan simultáneamente en T y satisfacen:*

a. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha I_\alpha(x_0,u_0) = 0$ $(\alpha \in R)$*;*

b. *Los multiplicadores $p_i(t)$ son continuos en T y en cada intervalo de continuidad de u_0 satisfacen*

$$\dot{x}^i = H_{p^i} = f^i, \quad \dot{p}_i = -H_{x^i}.$$

c. $H(t,x_0(t),u,p(t),\lambda,\lambda_0) \leq H(\tilde{x}_0(t),p(t),\lambda,\lambda_0)$ *para todo u en una vecindad de $u_0(t)$.*

d. *La función $H(\tilde{x}_0(t),p(t),\lambda,\lambda_0)$ es continua y satisface*

$$\frac{dH}{dt} = H_t$$

en los intervalos de continuidad de u_0.

e. $H_u(\tilde{x}_0(t),p(t),\lambda,\lambda_0) = 0 \quad \forall t \in T.$

Definición 2.2.2. *Denotamos por $\mathcal{E}$ a los extremos de (IP), es decir, al conjunto de $(x,u,p,\lambda) \in Z \times X \times \mathbb{R}^q$ tales que*

i. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha I_\alpha(x,u) = 0$ $(\alpha \in R)$*;*

ii. $\dot{p}(t) = -f_x^*(\tilde{x}(t))p(t) + L_x^*(\tilde{x}(t)) + \sum_1^q \lambda_\gamma L_{\gamma x}^*(\tilde{x}(t))$ $\forall t \in T$*;*

iii. $f_u^*(\tilde{x}(t))p(t) = L_u^*(\tilde{x}(t)) + \sum_1^q \lambda_\gamma L_{\gamma u}^*(\tilde{x}(t))$ $\forall t \in T$.

En [24] se define la normalidad de la siguiente manera.

Definición 2.2.3. *Diremos que (x_0,u_0) es fuertemente normal si existen $q+n$*

procesos $(y_\sigma, v_\sigma) \in Z$ *que satisfagan*

$$\dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t) \quad \forall t \in T, \quad y(t_0) = 0$$

y tales que el determinante

$$\Delta := \begin{vmatrix} I'_\gamma(x_0, u_0; y_\sigma, v_\sigma) \\ y^i_\sigma(t_1) \end{vmatrix} \qquad \begin{matrix} (\gamma = 1, \ldots, q;\ i = 1, \ldots, n) \\ (\sigma = 1, \ldots, q+n) \end{matrix}$$

sea diferente de cero donde, análogamente al caso anterior,

$$I'_\gamma(x_0, u_0; y, v) = \int_{t_0}^{t_1} \{L_{\gamma x}(\tilde{x}_0(t))y(t) + L_{\gamma u}(\tilde{x}_0(t))v(t)\}dt \quad ((y, v) \in Z)$$

denota la primera variación de I_γ *en* (x_0, u_0) *en la dirección de* (y, v).

En [24, pág. 275, Teorema 7.2] se prueba que, si (x_0, u_0) es solución local de (IP) fuertemente normal, es posible tomar $\lambda_0 = 1$ en el Teorema 2.2.1 y, en tal caso, los multiplicadores serán únicos, es decir, con las hipótesis anteriores existe un único par $(p, \lambda) \in X \times \mathbb{R}^q$ tal que $(x_0, u_0, p, \lambda) \in \mathcal{E}$.

Sea $\Lambda(L, x_0, u_0)$ el conjunto de multiplicadores (p, λ) tales que $(x_0, u_0, p, \lambda) \in \mathcal{E}$. Notemos que hay una diferencia sustancial entre $\Lambda(L, x_0, u_0)$ y $\Lambda(L, x_0)$ (del caso de cálculo de variaciones) pues este último, al igual que en el caso de dimensión finita, es un conjunto de vectores en $\mathbb{R}^q$. Sin embargo, $\Lambda(L, x_0, u_0)$ es un subconjunto de $X \times \mathbb{R}^q$, es decir, cada multiplicador se compone de un vector y de una función C^1 a trozos, por lo cual las técnicas utilizadas en las pruebas de los Teoremas 2.1.11 y 2.1.13 no pueden aplicarse a este problema.

Para remediar esto, consideramos las nociones de normalidad y de extremo, introducidas en [5, Definición 5.2.4]. Para ello, definimos al conjunto

$$Y_0 := \{(y, v) \in Z \mid \dot{y}(t) = A(t)y(t) + B(t)v(t)\ (t \in T),\ y(t_0) = 0\},$$

donde $A(t) = f_x(\tilde{x}_0(t))$, $B(t) = f_u(\tilde{x}_0(t))$, y a los funcionales lineales en Z,

$$F_\gamma(y, v) := I'_\gamma(x_0, u_0; y, v) \quad (\gamma = 1, \ldots, q),$$

$$F_{q+i}(y, v) := y^i(t_1) \quad (i = 1, \ldots, n).$$

Definición 2.2.4. *Diremos que* $(x_0, u_0) \in S$ *es normal con respecto a* S *si* $\eta = 0$ *es* *la única solución en $\mathbb{R}^{q+n}$ del sistema*

i. $\eta_\alpha \geq 0$ *y* $\eta_\alpha I_\alpha(x,u) = 0$ $(\alpha \in R)$;

ii. $\sum_1^{q+n} \eta_\gamma F_\gamma(y,v) = 0$ *para todo* $(y,v) \in Y_0$.

Definición 2.2.5. *Denotamos por* $\tilde{\mathcal{E}}$ *al conjunto de* $(x,u,\eta) \in S \times \mathbb{R}^{q+n}$ *tales que*

i. $\eta_\alpha \geq 0$ *y* $\eta_\alpha I_\alpha(x,u) = 0$ $(\alpha \in R)$;

ii. *Si* $\check{J}(\cdot,\cdot) := I'(x,u;\cdot,\cdot) + \sum_1^{q+n} \eta_\gamma F_\gamma(y,v) = 0$, *entonces* $\check{J}(y,v) = 0$ *para todo* $(y,v) \in Y_0$.

Nuestro siguiente teorema muestra que hay una biyección entre los conjuntos $\tilde{\mathcal{E}}$ y $\mathcal{E}$.

Teorema 2.2.6. *Para* $(x_0,u_0) \in S$ *se cumple lo siguiente:*

a. *Si* $(x_0,u_0,p,\lambda) \in \mathcal{E}$, *entonces* $(x_0,u_0,\lambda,-p^1(t_1),\ldots,-p^n(t_1)) \in \tilde{\mathcal{E}}$.

b. *Si* $(x_0,u_0,\eta) \in \tilde{\mathcal{E}}$, *entonces existe* $p \in X$ *tal que* $p^i(t_1) = -\eta_{q+i}$ $(i = 1,\ldots,n)$ *y satisface que* $(x_0,u_0,\eta_1,\ldots,\eta_q,p) \in \mathcal{E}$.

En particular, hay una biyección entre los conjuntos $\Lambda(L,x_0,u_0)$ *y*

$$\tilde{\Lambda}(L,x_0,u_0) := \{\eta \in \mathbb{R}^{q+n} \mid (x_0,u_0,\eta) \in \tilde{\mathcal{E}}\}.$$

Demostración.

(a): Sean $(x_0,u_0) \in S$ y $(p,\lambda) \in \Lambda(L,x_0,u_0)$. Si $(y,v) \in Y_0$, por 2.2.2(ii) tenemos que

$$y^*(t)\dot{p}(t) = -y^*(t)A^*(t)p(t) + y^*(t)L_x^*(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma y^*(t)L_{\gamma x}^*(\tilde{x}_0(t)) \ \forall t \in T \quad (2.2.2)$$

y, por 2.2.2(iii),

$$v^*(t)B^*(t)p(t) = v^*(t)L_u^*(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma v^*(t)L_{\gamma u}^*(\tilde{x}_0(t)) \ \forall t \in T. \quad (2.2.3)$$

Sumando las ecuaciones (2.2.2) y (2.2.3) e integrando sobre T obtenemos

$$\int_T \{y^*(t)\dot{p}(t) + y^*(t)A^*(t)p(t) + v^*(t)B^*(t)p(t)\}dt =$$

$$\int_T \Big(L_x(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma x}(\tilde{x}_0(t))\Big)y(t)dt + \int_T \Big(L_u(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma u}(\tilde{x}_0(t))\Big)v(t)dt. \quad (2.2.4)$$

Por la definición de Y_0, el lado izquierdo de (2.2.4) es igual a

$$\begin{aligned}\int_T \{y^*(t)\dot{p}(t) + \dot{y}^*(t)p(t)\}dt &= y^*(t_1)p(t_1) - y(t_0)p(t_0) \\ &= y^*(t_1)p(t_1) = \sum_1^n p^i(t_1)F_{q+i}(y,v),\end{aligned}$$

mientras que el lado derecho es igual a

$$I'(x_0,u_0;y,v) + \sum_1^q \lambda_\gamma I'_\gamma(x_0,u_0;y,v) = I'(x_0,u_0;y,v) + \sum_1^q \lambda_\gamma F_\gamma(y,v).$$

Así, $\eta := (\lambda, -p^1(t_1), \ldots, -p^n(t_1))$ satisface

$$I'(x_0,u_0;y,v) + \sum_1^{q+n} \eta_\gamma F_\gamma(y,v) = 0 \quad \text{para toda } (y,v) \in Y_0,$$

que es precisamente la condición 2.2.5(ii). Como 2.2.5(i) se sigue claramente de 2.2.2(i), tenemos que $\eta \in \tilde{\Lambda}(L, x_0, u_0)$.

(b): Sea $\eta \in \tilde{\Lambda}(L, x_0, u_0)$. Sean $\lambda := (\eta_1, \ldots, \eta_q)$ y p la única solución en T de la ecuación

$$\dot{p}(t) = -A(t)p(t) + L_x^*(\tilde{x}_0(t)) + \sum_1^q \eta_\gamma L_{\gamma x}^*(\tilde{x}_0(t)),$$

con condición final $p^i(t_1) = -\eta_{q+i}$ $(i = 1, \ldots, n)$.

Claramente $(p, \lambda) \in X \times \mathbb{R}^q$ y satisface las condiciones 2.2.2(i) y (ii). Para ver que 2.2.2(iii) también se cumple, consideremos $(y, v) \in Y_0$. Procediendo como en el inciso anterior, tenemos que

$$\sum_1^n p^i(t_1)y^i(t_1) = \int_T \{(L_x(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma x}(\tilde{x}_0(t)))y(t) + p^*(t)B(t)v(t)t\}dt. \quad (2.2.5)$$

Por la condición 2.2.5(ii)

$$\sum_1^n p^i(t_1)y^i(t_1) = I'(x_0,u_0;y,v) + \sum_1^q \lambda_\gamma I'(x_0,u_0;y,v). \quad (2.2.6)$$

Igualando las ecuaciones (2.2.5) y (2.2.6) obtenemos, para toda $(y, v) \in Y_0$,

$$\int_T p^*(t)B(t)v(t)dt = \int_T \{L_u(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma u}(\tilde{x}_0(t))\}v(t)dt. \quad (2.2.7)$$

Sin embargo, para toda $v \in \mathcal{U}$ existe $y \in X$ tal que $(y, v) \in Y_0$. Por tanto, la ecuación (2.2.7) es válida para toda $v \in \mathcal{U}$, implicando así la igualdad punto a punto, es decir,

$$B(t)p(t) = L_u^*(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t)) \ (t \in T).$$

Por lo tanto, $(p, \lambda) \in \Lambda(L, x_0, u_0)$. $\square$

Dados $(x_0, u_0) \in S$ y $(p, \lambda) \in \Lambda(L, x_0, u_0)$, definimos al conjunto de índices activos en (x_0, u_0) como

$$A(x_0, u_0) := \{\alpha \in R \mid I_\alpha(x_0, u_0) = 0\}$$

y definimos los conjuntos S_0 y S_1 como sigue:

$$S_0 \ (= S_0(x_0)) := \{(x, u) \in D \mid I_\gamma(x, u) = 0 \ (\gamma \in A(x_0, u_0) \cup Q)\},$$

$$S_1 \ (= S_1(\lambda)) := \{(x, u) \in S \mid I_\alpha(x, u) = 0 \ (\alpha \in R, \ \lambda_\alpha > 0)\}.$$

Aplicando la Definición 2.2.4 a estos conjuntos se tienen las siguientes definiciones.

Definición 2.2.7. *Diremos que $(x_0, u_0) \in S$ es normal con respecto a S_0 si $\eta = 0$ es la única solución del sistema*

i. $\eta_\alpha I_\alpha(x, u) = 0 \ (\alpha \in R)$;

ii. $\sum_1^{q+n} \eta_\gamma F_\gamma(y, v) = 0$ *para todo* $(y, v) \in Y_0$.

Definición 2.2.8. *Sea $(x_0, u_0) \in S$ para el cual existe $(p, \lambda) \in \Lambda(L, x_0, u_0)$. Diremos que (x_0, u_0) es normal con respecto a $S_1(\lambda)$ si $\kappa = 0$ es la única solución del sistema*

i. $\kappa_\alpha \geq 0$ *y* $\kappa_\alpha I_\alpha(x, u) = 0 \ (\alpha \in R, \ \lambda_\alpha = 0)$;

ii. $\sum_1^{q+n} \kappa_\gamma F_\gamma(y, v) = 0$ *para todo* $(y, v) \in Y_0$.

Con estas definiciones, la prueba del siguiente resultado, que caracteriza la unicidad de los multiplicadores de Lagrange, se obtiene aplicando el Teorema 2.2.6 y siguiendo la demostración del Teorema 2.1.11.

Teorema 2.2.9. *Sea $(x_0, u_0) \in S$ y supongamos que $(p, \lambda) \in \Lambda(L, x_0, u_0)$. Entonces son equivalentes:*

a. $\Lambda(L, x_0, u_0) = \{(p, \lambda)\}$.

b. (x_0, u_0) *es normal con respecto a* $S_1(\lambda)$.

Por otro lado, es fácil ver que la normalidad con respecto a S_0 es equivalente a la normalidad fuerte (ver detalles en [5]) y que ésta implica la normalidad con respecto a $S_1(\lambda)$ para cualquier $(p,\lambda) \in \Lambda(L, x_0, u_0)$. Por tanto, el Teorema 2.2.9 debilita la hipótesis del Teorema 7.2 de [24] para garantizar la unicidad del multiplicador (p,λ).

Finalmente, si $\tilde{\mathcal{F}}(x_0, u_0)$ denota al conjunto de funciones $L\colon T \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ de clase C^1 tales que (x_0, u_0) minimiza

$$I(x) = \int_{t_0}^{t_1} L(t, x(t), u(t))dt$$

sobre S, el Teorema 2.1.13 en este contexto establece lo siguiente.

Teorema 2.2.10. *Para* $(x_0, u_0) \in S$ *son equivalentes:*

a. (x_0, u_0) *es fuertemente normal.*

b. $\Lambda(L, x_0, u_0)$ *es unitario para toda* $L \in \tilde{\mathcal{F}}(x_0)$.

Demostración. (Note que (a) $\Rightarrow$ (b) es precisamente [24, Teorema 7.2]). Por el Teorema 2.2.6, (b) es equivalente a que $|\tilde{\Lambda}(L, x_0, u_0)| = 1$, de modo que la demostración es análoga a la del Teorema 2.1.13. □

Capítulo 3

El problema de control óptimo con restricciones puntuales

3.1. Planteamiento del problema y condiciones de primer orden

Es importante mencionar que los resultados expuestos en esta sección son válidos para problemas más generales, como el problema de Bolza con punto final variable (ver por ejemplo [31, 37]). Sin embargo, por simplicidad de exposición y de notación, analizaremos el problema de Lagrange con puntos inicial y final fijos.

Consideremos un intervalo $T := [t_0, t_1]$ en $\mathbb{R}$, dos puntos ξ_0, ξ_1 en $\mathbb{R}^n$, y funciones L y f que mapean $T \times \mathbb{R}^n \times \mathbb{R}^m$ a $\mathbb{R}$ y $\mathbb{R}^n$ respectivamente, y $\varphi = (\varphi_1, \ldots, \varphi_q)$ que mapea $\mathbb{R}^m$ a $\mathbb{R}^q$ $(q \leq m)$.

Nuevamente, denotemos por X al espacio de funciones C^1 a trozos mapeando T a $\mathbb{R}^n$, por $\mathcal{U}_k$ al espacio de funciones continuas a trozos mapeando T a $\mathbb{R}^k$ $(k \in \mathbb{N})$, y definamos los siguientes conjuntos: $Z := X \times \mathcal{U}_m$,

$$D := \{(x, u) \in Z \mid \dot{x}(t) = f(t, x(t), u(t)) \ \forall t \in T,\ x(t_0) = \xi_0,\ x(t_1) = \xi_1\},$$

$$S := \{(x, u) \in D \mid \forall t \in T,\ \varphi_\alpha(u(t)) \leq 0,\ \varphi_\beta(u(t)) = 0 \ (\alpha \in R,\ \beta \in Q)\},$$

donde $R = \{1, \ldots, r\}$, $Q = \{r + 1, \ldots, q\}$.

Consideremos el funcional $I \colon Z \to \mathbb{R}$ dado por

$$I(x, u) := \int_{t_0}^{t_1} L(t, x(t), u(t))dt \quad ((x, u) \in Z).$$

El problema que abordaremos, y que denotaremos como (P), consiste en minimizar I sobre S.

Llamaremos a los elementos de Z *procesos*, a los de S *procesos admisibles*, y diremos que un proceso (x, u) es *solución de* o *resuelve* (P) si (x, u) es admisible e $I(x, u) \leq I(y, v)$ para todo proceso admisible (y, v). Es decir, a diferencia de los capítulos previos y algunos trabajos como [31] y [46], consideraremos soluciones globales y no locales, sin perder generalidad, pues los resultados que enunciaremos sobre condiciones necesarias de primer y segundo orden son válidos si sustituimos $T \times \mathbb{R}^n \times \mathbb{R}^m$ por $T \times O \times V$ para cualesquiera conjuntos abiertos $O \subset \mathbb{R}^n$ y $V \subset \mathbb{R}^m$. Así, reduciendo estos conjuntos si fuera necesario, los resultados obtenidos serán válidos también para mínimos locales.

Recordemos que las funciones $x \in X$ se conocen como *estados* y las funciones $u \in \mathcal{U}_m$ como *controles*. Para simplificar la notación, como hemos hecho antes, dado $(x, u) \in Z$ representaremos a $(t, x(t), u(t))$ como $(\tilde{x}(t))$ (con $u(t)$ reemplazado por el límite lateral correspondiente si t es un punto de discontinuidad de u) y '$*$' denotará transpuesta.

Supondremos a lo largo de todo el capítulo que, si $G := (L, f)$, entonces $G(t, \cdot, \cdot)$ es C^1 (C^2 cuando aparezcan segundas derivadas) para todo $t \in T$, lo mismo que φ; $G(\cdot, x, u)$, $G_x(\cdot, x, u)$ y $G_u(\cdot, x, u)$ son continuas a trozos para todo $(x, u) \in \mathbb{R}^n \times \mathbb{R}^m$ y existe una función integrable $v\colon T \to \mathbb{R}$ tal que, para cualquier punto $(t, x, u) \in T \times \mathbb{R}^n \times \mathbb{R}^m$, se cumple

$$|G(t, x, u)| + |G_x(t, x, u)| + |G_u(t, x, u)| \leq v(t).$$

Supondremos también que la matriz de dimensión $q \times (m + r)$, dada por

$$\begin{pmatrix} \dfrac{\partial \varphi_i}{\partial u^k} & \delta_{i\alpha}\varphi_\alpha \end{pmatrix} \qquad (i = 1, \ldots, q;\ \alpha = 1, \ldots, r;\ k = 1, \ldots, m),$$

es de rango q en U (aquí $\delta_{\alpha\alpha} = 1$, $\delta_{\alpha\beta} = 0$ $(\alpha \neq \beta)$) y

$$U := \{u \in \mathbb{R}^m \mid \varphi_\alpha(u) \leq 0\ (\alpha \in R),\ \varphi_\beta(u) = 0\ (\beta \in Q)\}.$$

Esta condición es equivalente a que en cada punto u en U, la matriz

$$\begin{pmatrix} \dfrac{\partial \varphi_i}{\partial u^k} \end{pmatrix} \qquad (i = i_1, \ldots, i_p;\ k = 1, \ldots, m)$$

sea de rango p, donde $i_1, \ldots, i_p$ son los índices $i \in \{1, \ldots, q\}$ tales que $\varphi_i(u) = 0$. Los detalles de esta equivalencia se pueden ver en [19, Proposición 5.1]. Notemos que, a pesar de que el planteamiento parece muy similar, el tipo de restricciones que se consideran hacen a este problema radicalmente diferente a los casos anteriores, aun en el caso particular de cálculo de variaciones, es decir, cuando la dinámica está dada por $f(t, x, u) = u$.

El siguiente resultado, que establece condiciones necesarias de primer orden, corresponde a una versión del Principio Máximo de Pontryagin. Versiones más generales, como el Teorema 2.2.1, con sus demostraciones, pueden consultarse, por ejemplo, en [14, 24]. En este caso la función hamiltoniano H está dada por

$$H(t, x, u, p, \mu, \lambda) := \langle p, f(t, x, u)\rangle - \lambda L(t, x, u) - \langle \mu, \varphi(u)\rangle.$$

Teorema 3.1.1. *Si (x_0, u_0) es solución de (P), entonces existen $\lambda_0 \geq 0$, $p \in X$ y $\mu \in \mathcal{U}_q$ que no se anulan simultáneamente en T y satisfacen:*

a. *$\mu_\alpha(t) \geq 0$ y $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ t \in T)$;*

b. *$\dot{p}(t) = -H_x^*(\tilde{x}_0(t), p(t), \mu(t), \lambda_0)$ y $H_u(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) = 0$ en cada intervalo de continuidad de u_0;*

c. *$H(t, x_0(t), u, p(t), 0, \lambda_0) \leq H(\tilde{x}_0(t), p(t), 0, \lambda_0)$ para todo $(t, u) \in T \times U$.*

Al igual que en los problemas anteriores, basados en este teorema, se introducen los conceptos de *normalidad* (con respecto a S) y de *extremo*. El primero como la condición que nos permite asegurar que si (λ_0, p, μ) satisface las condiciones del Teorema 3.1.1, entonces λ_0, conocido como multiplicador del costo, no se anula, mientras que el segundo hace referencia a los procesos y los multiplicadores (p, μ) que cumplen dichas condiciones con $\lambda_0 = 1$.

Definición 3.1.2. *Decimos que $(x_0, u_0) \in S$ es un proceso normal con respecto a S o débilmente normal si, dado $(p, \mu) \in X \times \mathcal{U}_q$ que satisface*

i. *$\mu_\alpha(t) \geq 0$ y $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ \forall t \in T)$;*

ii. *$\dot{p}(t) = -f_x^*(\tilde{x}_0(t))p(t)$ $\quad [\, = -H_x^*(\tilde{x}_0(t), p(t), \mu(t), 0)\ ]$ $\forall t \in T$;*

iii. *$0 = f_u^*(\tilde{x}_0(t))p(t) - \varphi'^*(u_0(t))\mu(t)$ $\quad [\, = H_u^*(\tilde{x}_0(t), p(t), \mu(t), 0)\ ]$ $\forall t \in T$,*

se tiene que $p \equiv 0$ y, como consecuencia de la hipótesis en $\varphi'^*(u_0(t))$, $\mu \equiv 0$.

Definición 3.1.3. *Denotamos por $\mathcal{E}$ al conjunto de extremos, cuyos elementos son las funciones $(x, u, p, \mu) \in Z \times X \times \mathcal{U}_q$ tales que*

a. $\mu_\alpha(t) \geq 0$ *y* $\mu_\alpha(t)\varphi_\alpha(u(t)) = 0$ $(\alpha \in R,\ \forall t \in T)$*;*
b. $\dot{p}(t) = -f_x^*(\tilde{x}(t))p(t) + L_x^*(\tilde{x}(t))$ $\forall t \in T$*;*
c. $f_u^*(\tilde{x}(t))p(t) = L_u^*(\tilde{x}(t)) + \varphi'^*(u(t))\mu(t)$ $\forall t \in T$.

Para todo $(x_0, u_0) \in S$, denotamos por $\Lambda(L, x_0, u_0)$ al conjunto de multiplicadores $(p, \mu) \in X \times \mathcal{U}_q$ tales que $(x_0, u_0, p, \mu) \in \mathcal{E}$ y los llamaremos multiplicadores de Lagrange asociados a x_0. De las Definiciones 3.1.2 y 3.1.3 podemos deducir que, si (x_0, u_0) es una solución de (P), normal con respecto a S, entonces $\Lambda(L, x_0, u_0) \neq \emptyset$. Sin embargo, para asegurar la unicidad de estos multiplicadores es necesaria otra condición que, como mostraremos más adelante, es en general más restrictiva.

3.2. Conos críticos

Condiciones necesarias de segundo orden para (P) pueden consultarse, por ejemplo, en [14, 19, 24, 31]. Por lo general, dichas condiciones se expresan en términos de la segunda variación del hamiltoniano, una forma cuadrática que se define de la siguiente manera.

Definición 3.2.1. Segunda variación con respecto a H. *Para* $(x, u, p, \mu) \in Z \times X \times \mathcal{U}_q$ *sea*

$$J((x, u, p, \mu); (y, v)) := \int_{t_0}^{t_1} 2\Omega(t, y(t), v(t))dt \quad ((y, v) \in Z)$$

donde, para todo $(t, y, v) \in T \times \mathbb{R}^n \times \mathbb{R}^m$,

$$2\Omega(t, y, v) := -\big[\langle y, H_{xx}(t)y\rangle + 2\langle y, H_{xu}(t)v\rangle + \langle v, H_{uu}(t)v\rangle\big]$$

y $H(t)$ *denota* $H(\tilde{x}(t), p(t), \mu(t), 1)$.

En general, para una solución (x_0, u_0) de (P) tal que $(x_0, u_0, p, \mu) \in \mathcal{E}$, lo que se busca es asegurar la no negatividad de $J((x_0, u_0, p, \mu); (\cdot, \cdot))$ en ciertos subconjuntos de Z. El primero que se considera, pues generaliza de manera natural al conjunto $C_S(x_0)$ del caso en dimensión finita, es el conjunto de tangentes curvilíneos que definimos a continuación.

Definición 3.2.2. *Denotamos por* $C_S(x_0, u_0)$ *al conjunto de tangentes curvilíneos a* S *en* (x_0, u_0)*, es decir, los procesos* $(y, v) \in Z$ *para los cuales existen* $\delta > 0$ *y* $(x(\cdot, \epsilon), u(\cdot, \epsilon)) \in S$ $(0 \leq \epsilon < \delta)$ *que satisfacen*

a. $(x(t,0),u(t,0)) = (x_0(t),u_0(t))\ \forall t \in T$;
b. $(x_\epsilon(t,0),u_\epsilon(t,0)) = (y(t),v(t))\ \forall t \in T$;
c. $x(t,\cdot)$ *es de clase* C^2; $u(t,\cdot)$ *es de clase* C^2 *a trozos.*

Otro conjunto importante es el *cono tangente a S en x_0*, $\mathcal{T}_S(x_0,u_0)$ o cono de tangentes secuenciales. Para definirlo, recordemos que en Z consideramos la norma dada por (2.2.1), es decir,

$$\|z\| := \|(x,u)\| = \sup_{t\in T}(|x(t)| + |\dot{x}(t)|) + \sup_{t\in T}|u(t)|\,.$$

Definición 3.2.3. *Diremos que una sucesión $\{z_k = (x_k,u_k)\} \subset Z$ converge a $z_0 = (x_0,u_0)$ en la dirección $w = (y,v)$ si w es un proceso unitario ($\|w\| = 1$), $z_k \neq z_0$, y*

$$\lim_{k\to\infty} \|z_k - z_0\| = 0, \qquad \lim_{k\to\infty} \frac{z_k - z_0}{\|z_k - z_0\|} = w.$$

El cono tangente a S en z_0, denotado por $\mathcal{T}_S(z_0)$, es el cono cerrado determinado por los vectores unitarios w para los cuales existe una sucesión $\{z_k\} \subset S$ que converge a z_0 en la dirección w.

Equivalentemente, $\mathcal{T}_S(z_0)$ es el conjunto de todos los $w \in Z$ para los cuales existe una sucesión $\{z_k\} \subset S$ y una sucesión de números positivos $\{\epsilon_k\}$ tales que

$$\lim_{k\to\infty} \epsilon_k = 0, \qquad \lim_{k\to\infty} \frac{z_k - z_0}{\epsilon_k} = w.$$

Como se muestra en [5, pág. 63-68], la norma definida por (2.2.1) satisface que, si $\{z_k := (x_k,u_k)\}$ converge a $z_0 := (x_0,u_0)$ en la dirección $w := (y,v)$, se cumple que

$$\lim_{k\to\infty} \frac{I(z_k) - I(z_0)}{\|z_k - z_0\|} = I'(z_0;w). \tag{3.2.1}$$

$$\lim_{k\to\infty} \frac{I(z_k) - I(z_0) - I'(z_0;(z_k - z_0))}{\|z_k - z_0\|^2} = \frac{1}{2}I''(z_0;w). \tag{3.2.2}$$

Considerando lo anterior, podemos probar los siguientes resultados que establecen condiciones necesarias de segundo orden, sin imponer condiciones de normalidad, con un argumento similar al del Teorema 1.2.4.

Teorema 3.2.4. *Sea* $(x_0, u_0) \in S$ *para el cual existe* $(p, \mu) \in \Lambda(L, x_0, u_0)$*. Sea*

$$S_1 = S_1(\mu) := \{(x, u) \in S \mid \varphi_\alpha(u(t)) = 0 \ (\alpha \in R,\ \mu_\alpha(t) > 0,\ t \in T)\}.$$

Si (x_0, u_0) *resuelve (P), entonces*

$$J((x_0, u_0, p, \mu); (y, v)) \geq 0 \text{ para todo } (y, v) \in C_{S_1}(x_0, u_0).$$

Demostración. Definimos

$$K(x, u) := \langle p(t_1), \xi_1 \rangle - \langle p(t_0), \xi_0 \rangle + \int_{t_0}^{t_1} F(t, x(t), u(t))dt \quad ((x, u) \in Z)$$

donde, para todo $(t, x, u) \in T \times \mathbb{R}^n \times \mathbb{R}^m$,

$$F(t, x, u) := L(t, x, u) - \langle p(t), f(t, x, u) \rangle + \langle \mu(t), \varphi(u) \rangle - \langle \dot{p}(t), x \rangle.$$

Observemos que

$$F(t, x, u) = -H(t, x, u, p(t), \mu(t), 1) - \langle \dot{p}(t), x \rangle$$

y, si $(x, u) \in S$, entonces

$$K(x, u) = I(x, u) + \int_{t_0}^{t_1} \langle \mu(t), \varphi(u(t)) \rangle dt.$$

Notemos también que

$$S_1 = \{(x, u) \in S \mid K(x, u) = I(x, u)\}.$$

Como $(x_0, u_0) \in S_1$, (x_0, u_0) minimiza K en S_1.

Sea $(y, v) \in C_{S_1}(x_0, u_0)$ y sean $\delta > 0$ y $(x(\cdot, \epsilon), u(\cdot, \epsilon)) \in S_1$ $(0 \leq \epsilon < \delta)$ tales que

$$(x(t, 0), u(t, 0)) = (x_0(t), u_0(t)) \quad \text{y} \quad (x_\epsilon(t, 0), u_\epsilon(t, 0)) = (y(t), v(t)).$$

Entonces, $g(\epsilon) := K(x(\cdot,\epsilon), u(\cdot,\epsilon))$ $(0 \leq \epsilon < \delta)$ satisface

$$g(\epsilon) = I(x(\cdot,\epsilon), u(\cdot,\epsilon)) \geq I(x_0, u_0) = K(x_0, u_0) = g(0) \quad (0 \leq \epsilon < \delta).$$

Por otra parte,

$$F_x(\tilde{x}_0(t)) = -H_x(\tilde{x}_0(t), p(t), \mu(t), 1) - \dot{p}^*(t) = 0, \tag{3.2.3}$$

$$F_u(\tilde{x}_0(t)) = -H_u(\tilde{x}_0(t), p(t), \mu(t), 1) = 0. \tag{3.2.4}$$

Por lo tanto, $g'(0) = 0$ y $0 \leq g''(0) = J((x_0, u_0, p, \mu); (y, v))$. □

Teorema 3.2.5. *Sea* $(x_0, u_0) \in S$ *para el cual existe* $(p, \mu) \in \Lambda(L, x_0, u_0)$. *Si* (x_0, u_0) *resuelve (P), entonces*

$$J((x_0, u_0, p, \mu); (y, v)) \geq 0 \text{ para todo } (y, v) \in \mathcal{T}_{S_1}(x_0, u_0).$$

Demostración. Consideremos el funcional $K(x, u)$ definido en la prueba del Teorema 3.2.4. Observemos que

$$K''(x_0, u_0; y, v) = J(x_0, u_0; y, v) \quad \text{para todo } (y, v) \in Z.$$

Además, por las ecuaciones (3.2.3) y (3.2.4), se tiene que

$$K'(x_0, u_0; y, v) = 0 \quad \text{para todo } (y, v) \in Z.$$

Sea $(y, v) \in \mathcal{T}_{S_1}(x_0, u_0)$. Supongamos, sin pérdida de generalidad, que $\|(y, v)\| = 1$ y sea $\{z_k\}$ una sucesión en S_1 que converge a $z_0 := (x_0, u_0)$ en la dirección (y, v). Por la ecuación (3.2.2) y la minimalidad de z_0, tenemos que

$$0 \leq \lim_{k\to\infty} \frac{K(z_k) - K(z_0)}{\|z_k - z_0\|^2} = \frac{1}{2}K''(x_0, u_0; y, v) = \frac{1}{2}J(x_0, u_0; y, v).$$ □

Es importante mencionar que los conjuntos $C_S(z_0)$ y $\mathcal{T}_S(z_0)$ pueden resultar muy restrictivos y, en la mayoría de los casos, es muy complicado determinar si un proceso está o no en cualquiera de ellos. Por esta razón, es conveniente definir un conjunto análogo al cono linealizado $R_S(x_0)$ cuya pertenencia resulte más fácil de verificar. Para definirlo necesitaremos algunos conceptos previos.

Para todo $u \in \mathbb{R}^m$ denotamos como

$$I_a(u) := \{\alpha \in R \mid \varphi_\alpha(u) = 0\},$$

al conjunto de *índices activos en* $u \in \mathbb{R}^m$, y definimos el siguiente conjunto:

$$\tau(u) := \{h \in \mathbb{R}^m \mid \varphi'_\alpha(u)h \leq 0 \ (\alpha \in I_a(u)),\ \varphi'_\beta(u)h = 0 \ (\beta \in Q)\}.$$

Definición 3.2.6. *Para todo* $(x_0, u_0) \in S$ *sea*

$$\mathcal{L}(x_0, u_0) := \{(y, v) \in Z \mid \dot{y}(t) = A(t)y(t) + B(t)v(t) \ (\forall t \in T),\ y(t_0) = y(t_1) = 0\},$$

donde, al igual que en el capítulo anterior, $A(t) = f_x(\tilde{x}_0(t))$ *y* $B(t) = f_u(\tilde{x}_0(t))$, *y definimos al conjunto de procesos que satisfacen las restricciones tangenciales en* (x_0, u_0) *con respecto a* S *como*

$$\mathcal{R}_S(x_0, u_0) := \{(y, v) \in \mathcal{L}(x_0, u_0) \mid v(t) \in \tau(u_0(t)) \ (\forall t \in T)\}.$$

Nota 3.2.7. *Para todo* $(x_0, u_0) \in S$,

$$C_S(x_0, u_0) \subset \mathcal{R}_S(x_0, u_0) \quad y \quad \mathcal{T}_S(x_0, u_0) \subset \mathcal{R}_S(x_0, u_0).$$

Demostración. Sean $(y, v) \in C_S(x_0, u_0)$, $\delta > 0$ y $(x(\cdot, \epsilon), u(\cdot, \epsilon)) \in S$ $(0 \leq \epsilon < \delta)$ tales que

$$(x(t, 0), u(t, 0)) = (x_0(t), u_0(t)) \quad \text{y} \quad (x_\epsilon(t, 0), u_\epsilon(t, 0)) = (y(t), v(t)).$$

Entonces, para todo $0 \leq \epsilon < \delta$,

$$\dot{x}(t, \epsilon) = f(t, x(t, \epsilon), u(t, \epsilon)) \ (\forall t \in T), \quad x(t_0, \epsilon) = \xi_0,\ x(t_1, \epsilon) = \xi_1$$

Notemos que la propiedad (c) de $C_S(x_0, u_0)$ implica que las derivadas parciales cruzadas de $x(\cdot, \cdot)$ son iguales, por lo que

$$\dot{x}_\epsilon(t, 0) = \dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t),$$

$$y(t_0) = x_\epsilon(t_0, 0) = 0 \quad \text{y} \quad y(t_1) = x_\epsilon(t_1, 0) = 0$$

y, por lo tanto, $(y,v) \in \mathcal{L}(x_0,u_0)$. Además, para todo $(t,\epsilon) \in T \times [0,\delta)$,

$$\varphi_\alpha(u(t,\epsilon)) \leq 0 \ (\alpha \in R), \quad \varphi_\beta(u(t,\epsilon)) = 0 \ (\beta \in Q).$$

Para $i \in R \cup Q$ fijo, y $t \in T$, definimos $\gamma(\epsilon) := \varphi_i(u(t,\epsilon))$ de modo que, $\gamma'(0) = \varphi_i'(u_0(t))v(t)$. Así, si $i \in I_a(u_0(t))$ entonces $\gamma'(0) \leq 0$, y si $i \in Q$, $\gamma'(0) = 0$.

Para probar la segunda contención, hacemos una ligera modificación a la prueba de [5, Proposición 5.5.2]. Sean $(y,v) \in \mathcal{T}_S(x_0,u_0)$, $\{(x_k,u_k)\} \subset S$ y $\{\epsilon_k\} \subset \mathbb{R}^+$ tales que en la norma de Z se cumple que

$$\frac{(x_k - x_0, u_k - u_0)}{\epsilon_k} \to (y,v).$$

Notemos que esto implica que

$$\frac{(x_k - x_0)}{\epsilon_k} \to y, \quad \frac{(\dot{x}_k - \dot{x}_0)}{\epsilon_k} \to \dot{y}, \quad \frac{(u_k - u_0)}{\epsilon_k} \to v \quad \text{uniformemente en } T.$$

En particular,

$$y(t_0) = \lim_{k\to\infty} \frac{x_k(t_0) - x_0(t_0)}{\epsilon_k} = \lim_{k\to\infty} \frac{\xi_0 - \xi_0}{\epsilon_k} = 0.$$

Análogamente vemos que $y(t_1) = 0$.

Luego, para $t \in T$ fijo, por el desarrollo de Taylor de f alrededor de $(x_0(t),u_0(t))$, existe R_1 tal que

$$\begin{aligned} f(\tilde{x}_k(t)) = f(\tilde{x}_0(t)) &+ f_x(\tilde{x}_0(t))(x_k(t) - x_0(t)) \\ &+ f_u(\tilde{x}_0(t))(u_k(t) - u_0(t)) + R_1(x_k(t) - x_0(t), u_k(t) - u_0(t)), \end{aligned}$$

$$\lim_{(x,u)\to(x_0(t),u_0(t))} \frac{R_1(x - x_0(t), u - u_0(t))}{\|(x - x_0(t), u - u_0(t))\|} = 0.$$

Por lo tanto,

$$\begin{aligned} \frac{\dot{x}_k - \dot{x}_0}{\epsilon_k} &= \frac{f(t,x_k(t),u_k(t)) - f(t,x_0(t),u_0(t))}{\epsilon_k} \\ &= f_x(\tilde{x}_0(t))\frac{(x_k - x_0)}{\epsilon_k} + f_u(\tilde{x}_0(t))\frac{(u_k - u_0)}{\epsilon_k} + \frac{R_1(x - x_0(t), u - u_0(t))}{\epsilon_k} \end{aligned}$$

y, pasando al límite, obtenemos

$$\dot{y}(t) = A(t)y(t) + B(t)v(t).$$

Finalmente, si $\beta \in B$, tenemos que

$$\varphi'_\beta(u_0(t))v(t) = \lim_{k\to\infty} \frac{\varphi_\beta(u_k(t)) - \varphi_\beta(u_0(t))}{\epsilon_k} = 0$$

y, si $\alpha \in I_a(u_0(t))$, entonces

$$\varphi'_\alpha(u_0(t))v(t) = \lim_{k\to\infty} \frac{\varphi_\alpha(u_k(t))}{\epsilon_k} \leq 0. \qquad \square$$

Definición 3.2.8. *Sea $(x_0, u_0) \in S$. Diremos que (x_0, u_0) es un proceso regular con respecto a S si $C_S(x_0, u_0) = \mathcal{R}_S(x_0, u_0)$ y un proceso $\mathcal{T}$-regular con respecto a S si $\mathcal{T}_S(x_0, u_0) = \mathcal{R}_S(x_0, u_0)$.*

En contraste con los casos anteriores, la regularidad se define en términos del conjunto $C_S(x_0, u_0)$ de tangentes curvilíneos y no del cono tangente $\mathcal{T}_S(x_0, u_0)$. La razón de esta elección es que las pruebas de algunos resultados que relacionan las nociones de normalidad y regularidad en este contexto, están basadas en el teorema de la función implícita e involucran precisamente a los elementos de $C_S(x_0, u_0)$ (ver por ejemplo [19]).

Sin embargo, como se ha enfatizado, la prueba dada en [25] que muestra que normalidad con respecto a S implica regularidad con respecto a S, en el caso de dimensión finita, así como las generalizaciones al caso de restricciones isoperimétricas establecidas en [5] y [9], se basan en propiedades del cono $\mathcal{T}_S(x_0)$.

Además de $S_1(\mu)$, se definió un nuevo conjunto denotado por S_0 que considera únicamente las restricciones de igualdad y las de desigualdad correspondientes a índices activos de (x_0, u_0), así como su conjunto de direcciones críticas τ_0.

La normalidad con respecto a este conjunto se define de la misma manera que con respecto al conjunto S pero, al no existir desigualdades en consideración, no existe la restricción de signo que aparece en (i) de la Definición 3.2.1, por lo que la normalidad con respecto a S_0 es una condición más restrictiva que con respecto a S.

Para $(x_0, u_0) \in S$ fijo, definimos

$$S_0 := \{(x, u) \in D \mid \varphi_\alpha(u(t)) = 0 \ (\alpha \in I_a(u_0(t)) \cup Q, \ \forall t \in T)\},$$

$$\tau_0(u) := \{h \in \mathbb{R}^m \mid \varphi_i'(u)h = 0 \ (i \in I_a(u) \cup Q)\}.$$

Definición 3.2.9. *Decimos que* $(x_0, u_0) \in S$ *es un proceso normal con respecto a* S_0 *o fuertemente normal si, dado* $(p, \mu) \in X \times \mathcal{U}_q$ *que satisface*

i. $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0 \ (\alpha \in R, \ \forall t \in T)$;

ii. $\dot{p}(t) = -A^*(t)p(t) \quad [\, = -H_x^*(\tilde{x}_0(t), p(t), \mu(t), 0)\,] \ (\forall t \in T)$;

iii. $0 = B^*(t)p(t) - \varphi'^*(u_0(t))\mu(t) \quad [\, = H_u^*(\tilde{x}_0(t), p(t), \mu(t), 0)\,] \ (\forall t \in T)$,

se tiene que $p \equiv 0$ *y, por tanto,* $\mu \equiv 0$.

Aplicando la misma definición al conjunto $S_1(\mu)$ obtenemos.

Definición 3.2.10. *Si* $(x_0, u_0, p, \mu) \in \mathcal{E}$*, decimos que* (x_0, u_0) *es normal con respecto a* $S_1(\mu)$ *si, dado* $(z, \nu) \in X \times \mathcal{U}_q$ *tal que*

i. $\nu_\alpha(t) \geq 0$ *y* $\nu_\alpha(t)\varphi_\alpha(u_0(t)) = 0 \ (\alpha \in R, \ \mu_\alpha(t) = 0, \ \forall t \in T)$;

ii. $\dot{z}(t) = -A^*(t)z(t) \ (\forall t \in T)$;

iii. $0 = B^*(t)z(t) - \varphi'^*(u_0(t))\nu(t) \ (\forall t \in T)$,

se tiene que $z \equiv 0$ *y, en consecuencia,* $\nu \equiv 0$.

Para $u \in \mathbb{R}^m$ y $\mu \in \mathbb{R}^q$, definimos

$$\begin{aligned}\tau_1(u, \mu) := \{h \in \mathbb{R}^m \mid \varphi_\alpha'(u)h \leq 0 \ (\alpha \in I_a(u) \cap \Gamma_0(\mu)),\\ \varphi_\beta'(u)h = 0 \ (\beta \in \Gamma(\mu) \cup Q)\}\end{aligned}$$

donde $\Gamma_0(\mu) := \{\alpha \in R \mid \mu_\alpha = 0\}$ y $\Gamma(\mu) := \{\alpha \in R \mid \mu_\alpha \neq 0\}$.

Así, el conjunto de procesos que satisfacen las restricciones tangenciales con respecto a $S_1 = S_1(\mu)$ es

$$\mathcal{R}_{S_1(\mu)}(x_0, u_0) := \{(y, v) \in \mathcal{L}(x_0, u_0) \mid v(t) \in \tau_1(u_0(t), \mu(t)) \ (\forall t \in T)\}.$$

Nuestro siguiente resultado muestra que, pese a que el conjunto S_1 depende de μ, el conjunto $\mathcal{R}_{S_1(\mu)}(x_0, u_0)$ es independiente de dicho multiplicador, i.e., si $(p, \mu) \not\equiv (q, \nu) \in \Lambda(L, x_0, u_0)$, en general $S_1(\mu) \neq S_1(\nu)$ pero $\mathcal{R}_{S_1(\mu)}(x_0, u_0)$ y $\mathcal{R}_{S_1(\nu)}(x_0, u_0)$ coinciden.

Proposición 3.2.11. *Si* $(x_0, u_0) \in S$ *y* $(p, \mu) \in \Lambda(L, x_0, u_0)$*, entonces*

$$\mathcal{R}_{S_1(\mu)}(x_0, u_0) = \{(y, v) \in \mathcal{R}_S(x_0, u_0) \mid I'(x_0, u_0; y, v) = 0\}.$$

En particular, $\mathcal{R}_{S_1(\mu)}(x_0, u_0) = \mathcal{R}_{S_1(\nu)}(x_0, u_0)$ *para todo* $(q, \nu) \in \Lambda(L, x_0, u_0)$.

Demostración. Denotamos por

$$\mathcal{R}_{S_1}(x_0, u_0) := \{(y, v) \in \mathcal{R}_S(x_0, u_0) \mid I'(x_0, u_0; y, v) = 0\}.$$

Consideremos el funcional $K(x, u)$ definido en la prueba del Teorema 3.2.4,

$$K(x, u) := \langle p(t_1), \xi_1\rangle - \langle p(t_0), \xi_0\rangle + \int_{t_0}^{t_1} F(t, x(t), u(t))dt \quad ((x, u) \in Z)$$

donde

$$F(t, x, u) := L(t, x, u) - \langle p(t), f(t, x, u)\rangle + \langle \mu(t), \varphi(u)\rangle - \langle \dot{p}(t), x\rangle.$$

Sabemos que, si $(x, u) \in S$, entonces

$$K(x, u) = I(x, u) + \int_T \langle \mu(t), \varphi(u(t))\rangle dt.$$

Además, para todo $(y, v) \in Z$,

$$K'(x_0, u_0; y, v) = I'(x_0, u_0; y, v) + \int_T \mu^*(t)\varphi'(u_0(t))v(t)dt = 0.$$

Por lo tanto,

$$\int_T [(L_x(\tilde{x}_0(t)) - p^*(t)A(t) - \dot{p}^*(t))y(t) + \\ (L_u(\tilde{x}_0(t)) - p^*(t)B(t) + \mu^*(t)\varphi'(u_0(t)))v(t)]dt = 0.$$

Esto implica que

$$\int_T (L_x(\tilde{x}_0(t))y(t) + L_u(\tilde{x}_0(t))v(t))dt - \\ \int_T [p^*(t)(A(t)y(t) + B(t)v(t)) + \dot{p}^*(t))y(t)]dt + \int_T \mu^*(t)\varphi'(u_0(t))v(t)dt = 0.$$

Si $(y,v) \in \mathcal{L}(x_0,u_0)$, el lado izquierdo de la última ecuación se transforma en

$$\begin{aligned} I'(x_0,u_0;y,v) &- \int_T [p^*(t)\dot{y}(t) + \dot{p}^*(t)y(t)]dt + \int_T \mu^*(t)\varphi'(u_0(t))v(t)dt \\ &= I'(x_0,u_0;y,v) - [p^*(t)y(t)]_{t_0}^{t_1} + \int_T \mu^*(t)\varphi'(u_0(t))v(t)dt \\ &= I'(x_0,u_0;y,v) + \int_T \sum_{\gamma=1}^{q} \mu_\gamma(t)\varphi'_\gamma(u_0(t))v(t)dt. \end{aligned}$$

Si además $(y,v) \in \mathcal{R}_S(x_0,u_0)$, obtenemos

$$I'(x_0,u_0;y,v) + \int_T \sum_{\{\alpha|\mu_\alpha(t)>0\}} \mu_\alpha(t)\varphi'_\alpha(u_0(t))v(t)dt = 0.$$

Por lo tanto, $(y,v) \in \mathcal{R}_{S_1}(x_0,u_0)$ si y sólo si

$$\int_T \sum_{\{\alpha|\mu_\alpha(t)>0\}} \mu_\alpha(t)\varphi'_\alpha(u_0(t))v(t)dt = 0.$$

Pero $\mu_\alpha(t)\varphi'_\alpha(u_0(t))v(t) \leq 0$ $(\alpha \in R,\ \forall t \in T)$, lo que implica que la integral se anula si y sólo si $\varphi'_\alpha(u_0(t))v(t) = 0$ $(\alpha \in R,\ \mu_\alpha(t) > 0,\ \forall t \in T)$, i.e., si y sólo si $(y,v) \in \mathcal{R}_{S_1(\mu)}(x_0,u_0)$. □

3.3. Equivalencias de la condición de normalidad

Antes de analizar la relación entre las nociones de normalidad definidas previamente y los multiplicadores de Lagrange asociados a una solución de (P), así como su relación con las condiciones de segundo orden que se deben cumplir, daremos algunas caracterizaciones de esta condición.

3.3.1. τ-regularidad

Como se prueba en [39], la normalidad con respecto a los conjuntos S_0, S_1 y S, puede caracterizarse en términos de los conos τ_0, τ_1 y τ definidos arriba. Para probar dichas caracterizaciones necesitaremos el siguiente resultado auxiliar.

Lema 3.3.1. *Supongamos que, para algún $(x_0,u_0) \in S$ y $q \in X$, existe una función*

$\nu\colon T \to \mathbb{R}^q$ *tal que*

$$\nu_\alpha(t)\varphi_\alpha(u_0(t)) = 0 \ (\alpha \in R,\ t \in T), \quad q^*(t)B(t) = \nu^*(t)\varphi'(u_0(t)) \ (t \in T).$$

Entonces ν es continua a trozos en T, y continua en cada punto de continuidad de u_0.

Demostración. Sean $s \in T$, $v := u_0(s)$ y $J := \{i_1, \ldots, i_p\}$ el conjunto de índices $i \in R \cup Q$ tales que $\varphi_i(v) = 0$. Para $(i \in R,\ i \notin J)$ tendremos que $\varphi_i(v) < 0$ y, por la continuidad a trozos, existe un intervalo $[a, b]$ que contiene a s donde $\varphi_i(u_0(t)) < 0$, por tanto, $\nu_i(t) = 0$ $(i \in R,\ i \notin J,\ t \in [a, b])$.

Definimos $\hat{\nu} := (\nu_{i_1}, \ldots, \nu_{i_p})^*$ y $\hat{\varphi} := (\varphi_{i_1}, \ldots, \varphi_{i_p})^*$, de modo que $\hat{\varphi}'(v)$ tiene rango p y $|\hat{\varphi}'(v)\hat{\varphi}'^*(v)| \neq 0$. Reduciendo $[a, b]$ si fuera necesario para que esta desigualdad sea válida en todo el intervalo y u_0 sea continua, tendremos entonces que

$$q^*(t)B(t) = \hat{\nu}^*(t)\hat{\varphi}'(u_0(t)) \quad (t \in [a, b])$$

y, multiplicando por $\hat{\varphi}'^*(u_0(t))$, obtenemos

$$q^*(t)B(t)\hat{\varphi}'^*(u_0(t)) = \hat{\nu}^*(t)\hat{\varphi}'(u_0(t))\hat{\varphi}'^*(u_0(t)) \quad (t \in [a, b]).$$

Esta última ecuación tiene una solución continua $\hat{\nu}$, continua en el intervalo $[a, b]$. Explícitamente,

$$\hat{\nu}(t) = \phi^{-1}(t)\hat{\varphi}'(u_0(t))B^*(t)q(t)$$

donde $\phi(t) = \hat{\varphi}'(u_0(t))\hat{\varphi}'^*(u_0(t))$. □

Proposición 3.3.2. *Para $(x_0, u_0) \in S$, son equivalentes*

a. *(x_0, u_0) es normal con respecto a S_0.*

b. *$z \equiv 0$ es la única solución en X del sistema*

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h = 0 \ \text{ para todo } h \in \tau_0(u_0(t)) \ (\forall t \in T).$$

Si (x_0, u_0) satisface la condición (b), *se dice que es τ_0-regular.*

Demostración.

(b) $\Rightarrow$ (a): Supongamos que $(p, \mu) \in X \times \mathcal{U}_q$ satisface $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ para toda $\alpha \in R$ y para todo $t \in T$ y

$$\dot{p}(t) = -A^*(t)p(t), \quad B^*(t)p(t) = \varphi'^*(u_0(t))\mu(t).$$

Sea $h \in \tau_0(u_0(t))$. Entonces

$$p^*(t)B(t)h = \sum_{i=1}^{q} \mu_i(t)\varphi_i'(u_0(t))h = 0$$

implicando, por (b), que $p \equiv 0$.

(a) $\Rightarrow$ (b): Sea $z \in X$ tal que

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h = 0 \text{ para todo } h \in \tau_0(u_0(t)) \ (\forall t \in T). \tag{3.3.1}$$

Para cada $t \in T$, sea

$$\hat{\varphi} = (\varphi_{i_1}, \ldots, \varphi_{i_p}) \quad \text{donde} \quad I_a(u_0(t)) \cup Q = \{i_1, \ldots, i_p\}$$

y sea $\hat{\mu} = (\hat{\mu}_{i_1}, \ldots, \hat{\mu}_{i_p})$ dada por

$$\hat{\mu}(t) := \phi^{-1}(t)\hat{\varphi}'(u_0(t))B^*(t)z(t) \quad (\forall t \in T),$$

donde $\phi(t) = \hat{\varphi}'(u_0(t))\hat{\varphi}'^*(u_0(t))$. Definimos $\mu(t) = (\mu_1(t), \ldots, \mu_q(t))^*$ como

$$\mu_\alpha(t) := \begin{cases} \hat{\mu}_\alpha(t) & \text{si } \alpha \in I_a(u_0(t)) \cup Q \\ 0 & \text{en otro caso.} \end{cases}$$

Claramente, $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ t \in T)$ y $\hat{\mu}^*(t)\hat{\varphi}'(u_0(t)) = \mu^*(t)\varphi'(u_0(t))$ $(\forall t \in T)$. Sea

$$G(t) := I_{m\times m} - \hat{\varphi}^{*\prime}(u_0(t))\phi^{-1}(t)\hat{\varphi}'(u_0(t)).$$

Notemos que $\hat{\varphi}'(u_0(t))G(t) = 0$ $(\forall t \in T)$. Así, si $h_k(t)$ $(k = 1, \ldots, m)$ denota la k-ésima columna de $G(t)$, tenemos que

$$\varphi_{i_j}'(u_0(t))h_k(t) = 0 \quad (j = 1, \ldots, p,\ k = 1, \ldots, m),$$

es decir, $h_k(t) \in \tau_0(u_0(t))$ y, por (3.3.1), $z^*(t)B(t)h_k(t) = 0$. Obtenemos entonces que

$$z^*(t)B(t) - \mu^*(t)\varphi'(u_0(t)) = z^*(t)B(t)G(t) = 0 \tag{3.3.2}$$

y, por el Lema 3.3.1, $\mu \in \mathcal{U}_q$. Por tanto, (a) implica $z \equiv 0$. $\square$

Para conjuntos definidos por igualdades y desigualdades se tienen las siguientes equivalencias.

Proposición 3.3.3. *Para $(x_0, u_0) \in S$, son equivalentes*

a. (x_0, u_0) *es normal con respecto a S.*

b. $z \equiv 0$ *es la única solución en X del sistema*

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h \leq 0 \ \text{ para todo } h \in \tau(u_0(t)) \ (\forall t \in T).$$

Si (x_0, u_0) satisface la condición (b)*, se dice que es τ-regular.*

Demostración.

(b) $\Rightarrow$ (a): Supongamos que $(p, \mu) \in X \times \mathcal{U}_q$ satisface

$$\mu_\alpha(t) \geq 0, \quad \mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0 \quad (\alpha \in R, \ t \in T),$$

$$\dot{p}(t) = -A^*(t)p(t), \quad B^*(t)p(t) = \varphi'^*(u_0(t))\mu(t).$$

Sea $h \in \tau(u_0(t))$. Entonces

$$p^*(t)B(t)h = \sum_{i=1}^{q} \mu_i(t)\varphi_i'(u_0(t))h \leq 0$$

implicando, por (b), que $p \equiv 0$.

(a) $\Rightarrow$ (b): Sea $z \in X$ tal que

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h \leq 0 \ \forall h \in \tau(u_0(t)) \quad (\forall t \in T). \tag{3.3.3}$$

Procediendo como en (a) $\Rightarrow$ (b) de la prueba de la Proposición 3.3.2, vemos que $h_k(t) \in \tau_0(u_0(t))$, pero $\tau_0(u_0(t)) \subseteq \tau(u_0(t))$. Entonces, por (3.3.3),

$$z^*(t)B(t)h_k(t) \leq 0 \quad (k = 1, \ldots, m).$$

Sin embargo, como $-h_k(t) \in \tau_0(u_0(t))$, la desigualdad anterior es, de hecho, una igualdad. Por lo tanto, de la ecuación (3.3.2) obtenemos que

$$z^*(t)B(t) = \mu^*(t)\varphi'(u_0(t)) \tag{3.3.4}$$

y, nuevamente por Lema 3.3.1, bastará mostrar que $\mu_\alpha(t) \geq 0$ $(\alpha \in R,\ t \in T)$ para obtener el resultado.

Para ver que, en efecto, μ satisface esta condición, consideremos la $p \times m$ matriz definida por $\hat{C}(t) := \phi(t)^{-1}\hat{\varphi}'(u_0(t))$. Observemos que

$$\hat{C}(t)\hat{\varphi}'^*(u_0(t)) = I_{p\times p} = \hat{\varphi}'(u_0(t))\hat{C}^*(t). \tag{3.3.5}$$

Por lo tanto, si $i_j \in I_a(u_0(t))$, las columnas $\hat{c}_j(t)$ de la transpuesta de la matriz $\hat{C}(t)$ satisfacen lo siguiente:

$$\varphi'_k(u_0(t))(-\hat{c}_j(t)) = \begin{cases} -1 & \text{si} \quad k = i_j \\ 0 & \text{si} \quad k \neq i_j \end{cases}$$

implicando que $-\hat{c}_j(t) \in \tau(u_0(t))$. De lo anterior obtenemos $z^*(t)B(t)\hat{c}_j(t) \geq 0$ $(\forall t \in T)$, pero, por (3.3.4) y (3.3.5) $\hat{\mu}^*(t) = z^*(t)B(t)\hat{C}^*(t)$. Así, si $\alpha \in I_a(u_0(t))$, $\hat{\mu}_\alpha(t) \geq 0$. □

Con una prueba análoga a la de la proposición anterior se obtiene el siguiente resultado.

Proposición 3.3.4. *Si* $(x_0, u_0) \in S$ *y* $\mu \in \mathcal{U}_q$ *satisface* $\mu_\alpha(t) > 0$ *y* $\mu_\alpha(t)\varphi_\alpha(u_0(t))$ $(\alpha \in R,\ t \in T)$*, son equivalentes:*

a. (x_0, u_0) *es normal con respecto a* $S_1(\mu)$*.*

b. $z \equiv 0$ *es la única solución en* X *del sistema*

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h \leq 0 \ \textit{ para todo } h \in \tau_1(u_0(t), \mu(t)) \quad (\forall t \in T).$$

Si (x_0, u_0) *satisface la condición* (b)*, se dice que es* τ_1*-regular.*

El siguiente resultado [39, Proposición 3.6] muestra que normalidad fuerte implica normalidad con respecto a $S_1(\mu)$ (μ como en la proposición anterior) y ésta, a su vez, implica normalidad débil.

Proposición 3.3.5. *Sean* $(x_0, u_0) \in S$ *y* $\mu \in \mathcal{U}_q$ *tal que*

$$\mu_\alpha(t) \geq 0, \quad \mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0 \ (\alpha \in R, \ \forall t \in T)$$

y considere los siguientes enunciados:

i. (x_0, u_0) *es* τ_0*-regular.*

ii. (x_0, u_0) *es* τ_1*-regular.*

iii. (x_0, u_0) *es* τ*-regular.*

Entonces (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii).

Demostración.

(i) $\Rightarrow$ (ii): Sea $z \in X$ tal que

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h \leq 0 \ \forall h \in \tau_1(u_0(t), \mu(t)) \quad (\forall t \in T).$$

Claramente $\tau_0(u_0(t)) \subseteq \tau_1(u_0(t), \mu(t))$ pues, si $h \in \tau_0(u_0(t))$, se tiene que

$$\varphi_i'(u_0(t))h = 0 \text{ para todo } i \in I_a(u_0(t)).$$

Entonces, por hipótesis, $z^*(t)B(t)h \leq 0$. Sin embargo, dado que $\tau_0(u_0(t))$ es un subespacio, tenemos que $z^*(t)B(t)h = 0$ implicando, por (i), que $z \equiv 0$.

(ii) $\Rightarrow$ (iii): Sea $z \in X$ tal que

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h \leq 0 \ \forall h \in \tau(u_0(t)) \quad (\forall t \in T).$$

Como en el caso anterior, es fácil ver que $\tau_1(u_0(t), \mu(t)) \subseteq \tau(u_0(t))$. De modo que, en particular, $z^*(t)B(t)h \leq 0 \ \forall h \in \tau_1(u_0(t), \mu(t))$. Por lo tanto, (ii) implica que $z \equiv 0$. □

3.3.2. Controlabilidad

En esta sección mostramos que la normalidad con respecto a un conjunto S, definido por igualdades y desigualdades, es también equivalente a la controlabilidad de un sistema lineal en T.

Dadas las funciones matriciales $C\colon T \to \mathbb{R}^{n \times n}$, $D\colon T \to \mathbb{R}^{n \times m}$ y un subconjunto

$\mathcal{V} \subset \mathcal{U}_m$ tal que $0 \in \mathcal{V}$, considere el siguiente sistema lineal

$$\begin{aligned} \dot{y}(t) &= C(t)y(t) + D(t)v(t) \\ v &\in \mathcal{V}. \end{aligned} \tag{3.3.6}$$

Definición 3.3.6. *Se dice que el sistema (3.3.6) es controlable sobre $\mathcal{V}$ si, para todo $w \in \mathbb{R}^n$, existe $v \in \mathcal{V}$ tal que la respuesta a*

$$\dot{y}(t) = C(t)y(t) + D(t)v(t), \quad y(t_0) = 0$$

satisface $y(t_1) = w$.

Teorema 3.3.7. *Un proceso $(x_0, u_0) \in S$ es normal con respecto a S si y sólo si la ecuación*

$$\dot{y}(t) = A(t)y(t) + B(t)v(t) \tag{3.3.7}$$

es controlable sobre

$$\mathcal{V} = \{v \in \mathcal{U}_m \mid v(t) \in \tau(u_0(t)) \ (\forall t \in T)\}.$$

Demostración. Sea $X(t)$ la matriz fundamental del sistema $\dot{z}(t) = A(t)z(t)$ (la parte homogénea de 3.3.7), principal en t_0, es decir $X(t_0) = I_{n\times n}$. Para $v \in \mathcal{U}_m$, la solución de

$$\dot{y}(t) = A(t)y(t) + B(t)v(t), \quad y(t_0) = 0$$

está dada por

$$y(t) = \int_{t_0}^{t} X(t_1)X^{-1}(s)B(s)v(s)ds \quad (t \in T).$$

De modo que la controlabilidad del sistema sobre $\mathcal{V}$ es equivalente a la suprayectividad del mapeo $\kappa\colon \mathcal{V} \to \mathbb{R}^n$ dado por:

$$\kappa(v) = \int_{t_0}^{t_1} X(t_1)X^{-1}(s)B(s)v(s)ds.$$

Notemos que $\kappa(\mathcal{V})$ es un subconjunto con estructura de cono convexo que contiene al 0, pues κ es lineal y $\mathcal{V}$ es convexo y contiene a la función cero. Así, si $\kappa(\mathcal{V}) \neq \mathbb{R}^n$, entonces 0 es un punto de su frontera y, por [23, Teorema 2.2.4], existe un hiperplano H que soporta a $\kappa(\mathcal{V})$ en 0, es decir, existe $\gamma \neq 0$ tal que $\displaystyle\inf_{w\in\kappa(\mathcal{V})} \langle \gamma, w\rangle = 0$, por lo tanto,

$$\forall v \in \mathcal{V}, \quad 0 \leq \langle \gamma, \kappa(v) \rangle = \left\langle \gamma, \int_T X(t_1)X^{-1}(s)B(s)v(s)ds \right\rangle$$

$$= \int_T \langle B^*(s)X^{*-1}(s)X^*(t_1)\gamma, v(s) \rangle \, ds.$$

Observemos que $y(s) := X^{*-1}(s)X^*(t_1)\gamma$, para $s \in T$, es una solución no nula de la ecuación $\dot{y}(s) = -A^*(s)y(s)$, pues ésta es la ecuación adjunta de la parte homogénea de (3.3.7) y, como la desigualdad anterior se debe cumplir para toda $v \in \mathcal{V}$, se deberá satisfacer punto a punto, es decir,

$$\langle B^*(t)y(t), h \rangle \geq 0 \ \forall h \in \tau(u_0(t)) \quad (\forall t \in T).$$

Por lo tanto, $-y(t)$ es una solución no nula del sistema

$$\dot{z}(t) = -A^*(t)z(t), \quad \langle B^*(t)z(t), h \rangle \leq 0 \ \forall h \in \tau(u_0(t)) \quad (\forall t \in T)$$

lo que implica, por la Proposición 3.3.2, que (x_0, u_0) no es débilmente normal. □

Análogamente, se prueban los siguientes resultados.

Corolario 3.3.8. *Un proceso $(x_0, u_0) \in S$ es normal con respecto a S_0 si y sólo si la ecuación*

$$\dot{y}(t) = A(t)y(t) + B(t)v(t) \tag{3.3.8}$$

es controlable sobre

$$\mathcal{V}_0 = \{v \in \mathcal{V} \mid v(t) \in \tau_0(u_0(t)) \ (\forall t \in T)\}.$$

Corolario 3.3.9. *Si $(x_0, u_0) \in S$ y $(p, \mu) \in \Lambda(L, x_0, u_0)$, entonces (x_0, u_0) es normal con respecto a $S_1(\mu)$ si y sólo la ecuación*

$$\dot{y}(t) = A(t)y(t) + B(t)v(t) \tag{3.3.9}$$

es controlable sobre

$$\mathcal{V}_1 = \{v \in \mathcal{V} \mid v(t) \in \tau_1(u_0(t), \mu(t)) \ (\forall t \in T)\}.$$

3.3.3. Propiedad

En los problemas del capítulo anterior la condición de normalidad resultó, como en dimensión finita, ser equivalente a una condición del tipo Mangasarian-Fromovitz. En esta sección mostraremos que, bajo la hipótesis adicional de que la función conjunto valuada $I_a(u_0(t))$ sea constante a trozos en T, la normalidad (con respecto a S) en este contexto también es equivalente a una condición de ese tipo. Para ello, aplicaremos la Definición 3.1.2 a dos subconjuntos más de Z.

Definición 3.3.10. *Decimos que* $(x_0, u_0) \in Z$ *es normal con respecto a* D *si, dado* $p \in X$ *que satisface*

i. $\dot{p}(t) = -A^*(t)p(t) \quad (\forall t \in T)$;

ii. $B^*(t)p(t) = 0 \quad (\forall t \in T)$,

se tiene que $p \equiv 0$.

La definición anterior es [19, Definición 2.3], donde se considera un problema que no está sujeto a restricciones de igualdad ni desigualdad, es decir, se busca minimizar al funcional $I(x, u)$ sobre el conjunto D. En el mismo texto, se considera también la normalidad de un proceso con respecto a un conjunto definido únicamente por restricciones de igualdad, [Definición 4.4], como el que definimos a continuación:

$$S_e := \{(x, u) \in D \mid \forall t \in T, \varphi_\beta(u(t)) = 0 \ (\beta \in Q)\}.$$

Definición 3.3.11. *Decimos que* $(x_0, u_0) \in Z$ *es normal con respecto a* S_e *si, dado* $(p, \mu) \in X \times \mathcal{U}_{q-r}$ *que satisface*

i. $\dot{p}(t) = -A^*(t)p(t) \quad (\forall t \in T)$;

ii. $0 = B^*(t)p(t) - \bar{\varphi}'^*(u_0(t))\mu(t) \quad (\forall t \in T)$,

(donde $\bar{\varphi} := (\varphi_{r+1}, \ldots, \varphi_q)$*) se tiene que* $p \equiv 0$ *y, por tanto,* $\mu \equiv 0$.

Considerando lo anterior, como en [9], definimos el concepto de *propiedad* de la siguiente manera.

Definición 3.3.12. *Diremos que* $(x_0, u_0) \in S$ *es propio con respecto a* S*, si satisface:*

i. *Si* $Q \neq \emptyset$*, entonces* (x_0, u_0) *es normal con respecto a* S_e*. Si, por el contrario, no hay restricciones de igualdad, entonces* (x_0, u_0) *es normal con respecto a* D*.*

ii. *Existe* $(y, v) \in \mathcal{R}_S(x_0, u_0)$ *tal que, para* $\alpha \in I_a(u_0(t))$ $(\forall t \in T)$*, se satisface* $\varphi'_\alpha(u_0(t))v(t) < 0$*.*

Teorema 3.3.13. *Sea* $(x_0, u_0) \in S$ *tal que* $I_a(u_0(t))$ *es constante a trozos. Entonces* (x_0, u_0) *es débilmente normal si y sólo si es propio con respecto a* S.

Demostración. Si (x_0, u_0) es débilmente normal, satisface la condición de normalidad con respecto a D o a S_e trivialmente. Ahora, si para cada $t \in T$ definimos

$$\hat{\varphi} := (\varphi_{i_1}, \ldots, \varphi_{i_p}) \quad \text{donde } i_j \in I_a(u_0(t)) \cup Q$$

entonces, para $i_j \in I_a(u_0(t))$, las columnas $\hat{c}_j(t)$ de la transpuesta de la matriz

$$\hat{C}(t) := (\hat{\varphi}'(u_0(t))\hat{\varphi}'(u_0(t))^*)^{-1}\hat{\varphi}'(u_0(t))$$

satisfacen

$$\varphi_k'(u_0(t))(-\hat{c}_j(t)) = \begin{cases} -1 & \text{si} \quad k = i_j \\ 0 & \text{si} \quad k \neq i_j. \end{cases}$$

Además, por la constancia a trozos, para $\alpha \in R$, la función $c_\alpha \colon T \to \mathbb{R}^m$ dada por

$$c_\alpha(t) := \begin{cases} \hat{c}_\alpha(t) & \alpha \in I_a(u_0(t)) \\ 0 & \text{en otro caso} \end{cases}$$

es también continua a trozos.

Sea $\hat{y}(t)$ la única respuesta al sistema

$$\dot{y}(t) = A(t)y(t) + B(t)\left(\sum_{\alpha \in R} c_\alpha(t)\right), \quad y(t_0) = 0.$$

Por el Teorema 3.3.7, sabemos que existe $v_c(\cdot) \in \mathcal{V}$ tal que si $y_c(\cdot)$ es la solución de

$$\dot{y}(t) = A(t)y(t) + B(t)v_c(t), \quad y(t_0) = 0$$

entonces se cumple que $y_c(t_1) = \hat{y}(t_1)$.

Recordemos que

$$\mathcal{V} = \{v \in \mathcal{U}_m \mid v(t) \in \tau(u_0(t)) \ (\forall t \in T)\}.$$

Por lo tanto, $v(t) := v_c(t) - \sum_\alpha c_\alpha(t) \in \mathcal{V}$ y satisface que $\varphi_\alpha'(u_0(t))v(t) \leq -1 < 0$ para toda $\alpha \in I_a(u_0(t))$, para todo $t \in T$.

Más aún, si $X(t)$ denota la matriz fundamental del sistema $\dot{z}(t) = A(t)z(t)$ tal que $X(t_0) = I_{n\times n}$ y $y(\cdot)$ es la solución de

$$\dot{y}(t) = A(t)y(t) + B(t)v(t), \quad y(t_0) = 0,$$

entonces

$$y(t_1) = y_c(t_1) - \hat{y}(t_1) = 0,$$

por tanto, $(y, v) \in \mathcal{R}_S(x_0, u_0)$, lo que muestra que la normalidad implica una condición del tipo Mangasarian-Fromovitz.

Recíprocamente, consideremos $(p, \mu) \in X \times \mathcal{U}_q$ una solución del sistema 3.1.2, es decir,

$$\mu_\alpha(t) \geq 0, \quad \mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0 \quad (\alpha \in R, \ \forall t \in T),$$

$$\dot{p}(t) = -A^*(t)p(t), \quad p^*(t)B(t) = \mu^*(t)\varphi'(u_0(t)) \quad (\forall t \in T)$$

y $(y, v) \in \mathcal{R}_S(x_0, u_0)$ tal que $\varphi'_\alpha(u_0(t))v(t) < 0$ para $\alpha \in I_a(u_0(t))$ $(\forall t \in T)$. Entonces

$$\int_T \mu^*(t)\varphi'(u_0(t))v(t)dt = \int_T p^*(t)B(t)v(t)dt = \int_T p^*(t)\left[\dot{y}(t) - A(t)y(t)\right]dt$$

$$= \int_T \left[p^*(t)\dot{y}(t) + \dot{p}^*(t)y(t)\right]dt = p^*(t_1)y(t_1) - p^*(t_0)y(t_0) = 0.$$

Por tanto, $\mu_\alpha(t) \equiv 0$ $\forall \alpha \in R$. Así, si $Q \neq \emptyset$, (p, μ) es una solución de 3.3.11(i)-(iii) y si $Q = \emptyset$, (p, μ) satisface 3.3.10(i)-(iii). Por lo tanto, de la normalidad con respecto a S_e o a D, se sigue que $p \equiv 0$ y $\mu \equiv 0$. $\square$

Extender esta caracterización a $S_1(\mu)$ (con $(p, \mu) \in \Lambda(L, x_0, u_0)$) no resulta inmediato como en los casos anteriores pues, además de la función de índices activos, deberemos pedir que las funciones

$$\Gamma_0(t) := \{\alpha \in R \mid \mu_\alpha(t) = 0\} \quad \text{y} \quad \Gamma_+(t) := \{\alpha \in R \mid \mu_\alpha(t) > 0\},$$

definidas en el intervalo T sean también constantes a trozos para poder construir una variación $(y, v) \in \mathcal{R}_{S_1}(x_0, u_0)$ que satisfaga las desigualdades de manera estricta. Más importante aún, a menos de que $\Gamma_+(t)$ sea constante en T, no es posible adaptar la condición 3.3.12(i), pues no podríamos asociar a $S_1(\mu)$ un conjunto S_{1e} definido por igualdades (como S_e a S).

Hemos encontrado así una diferencia fundamental entre este problema y los anteriores, pues recordemos que esta caracterización de la normalidad fue crucial en las pruebas de [5, 9, 25] que muestran que normalidad implica regularidad, obteniendo así condiciones necesarias de segundo orden para optimalidad en el cono R_{S_1}. Asimismo, la prueba de la equivalencia entre normalidad con respecto a $S_1(\lambda)$ y unicidad del multiplicador λ que se da en [23, Teorema 3.10.4] depende explícitamente de la condición estricta de Mangasarian-Fromovitz.

Considerando lo anterior, dada una solución (x_0, u_0) de (P) y un multiplicador de Lagrange $(p, \mu) \in \Lambda(L, x_0, u_0)$, es natural preguntarse ¿cuál es la relación entre la condición de normalidad con respecto a $S_1(\mu)$, la unicidad de (p, μ) y el signo de $J((x_0, u_0, p, \mu); (\cdot, \cdot))$ en el cono $\mathcal{R}_{S_1}(x_0, u_0)$? Intentamos responder estas interrogantes en las siguientes secciones.

3.4. Normalidad y unicidad de multiplicadores de Lagrange

En esta sección, hacemos un análisis de la relación entre los distintos tipos de normalidad definidos en la sección previa y la unicidad de los multiplicadores de Lagrange para el problema de control óptimo pues, como mencionamos en la introducción, el tema de la unicidad ha recibido poca atención.

Además del trabajo de Malanowski [32], donde se dan condiciones de suficiencia para la unicidad, no parece haber en la literatura textos donde se busque caracterizarla. En algunas referencias como [19] o [24], se prueba la siguiente condición suficiente para unicidad, como parte de un resultado referente a condiciones de segundo orden.

Nota 3.4.1. *Si (x_0, u_0) es una solución de (P) fuertemente normal, entonces existe un único par $(p, \mu) \in X \times \mathcal{U}_q$ tal que $(p, \mu) \in \Lambda(L, x_0, u_0)$.*

Demostración. Por Proposición 3.3.5 sabemos que la normalidad fuerte implica normalidad débil, por tanto, existe $(p, \mu) \in \Lambda(L, x_0, u_0)$. Supongamos que (q, ν) también pertenece a $\Lambda(L, x_0, u_0)$. Tenemos entonces

i. $(\mu_\alpha - \nu_\alpha)(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ \forall t \in T)$;

ii. $[\dot{q} - \dot{p}]\,(t) = -A^*(t)\,[q - p]\,(t) \quad (\forall t \in T)$;

iii. $B^*(t)\,[q - p]\,(t) = {\varphi'}^*(u_0(t))\,[\mu - \nu]\,(t) \quad (\forall t \in T)$,

implicando, por normalidad fuerte, que $p \equiv q$ y $\mu \equiv \nu$. □

Sin embargo, esta hipótesis puede debilitarse y, dado $(p, \mu) \in \Lambda(L, x_0, u_0)$, la unicidad se sigue si el proceso es normal con respecto a $S_1(\mu)$, como mostramos a continuación.

3.4.1. Los resultados de Kyparisis y Wachsmuth en control óptimo

Teorema 3.4.2. *Sea* $(x_0, u_0) \in S$ *y suponga que existe* $(p, \mu) \in X \times \mathcal{U}_q$ *tal que* $(p, \mu) \in \Lambda(L, x_0, u_0)$. *Si* (x_0, u_0) *es normal con respecto a* $S_1(\mu)$, *entonces* $\Lambda(L, x_0, u_0) = \{(p, \mu)\}$.

Demostración. Supongamos que $(\bar{p}, \bar{\mu}) \in \Lambda(L, x_0, u_0)$ y sea $(q, \nu) := (\bar{p}-p, \bar{\mu}-\mu)$. Mostraremos que (q, ν) satisface las condiciones de la Definición 3.2.10. Si $\alpha \in R$ y $\mu_\alpha(t) = 0$, entonces

$$\nu_\alpha(t) = \bar{\mu}_\alpha(t) \geq 0 \quad \text{y} \quad \nu_\alpha(t)\varphi_\alpha(u_0(t)) = \bar{\mu}_\alpha(t)\varphi_\alpha(u_0(t))$$

y entonces 3.2.10(i) se satisface. Las otras dos condiciones, 3.2.10(ii) y 3.2.10(iii) se siguen de

$$\dot{q}(t) = \dot{\bar{p}}(t) - \dot{p}(t) = -A^*(t)q(t) \quad (\forall t \in T),$$

$$B^*(t)q(t) = B^*(t)(\bar{p}(t) - p(t)) = \varphi'^*(u_0(t))\nu(t) \quad (\forall t \in T).$$

Como (x_0, u_0) es normal con respecto a $S_1(\mu)$, $(q, \nu) \equiv (0, 0)$. □

El teorema anterior corresponde, en los problemas previos, al resultado de suficiencia para unicidad de los Teoremas 1.3.3, 2.1.11 y 2.2.9. El recíproco, sin embargo, no necesariamente es cierto en este contexto pues, si (x_0, u_0) no es normal con respecto a $S_1(\mu)$, existe $(q, \nu) \in X \times \mathcal{U}_q$ con $q \not\equiv 0$ satisfaciendo 3.2.10(i)-(iii) pero, contrario a los casos anteriores, no podemos garantizar que ν pueda ser elegido de forma que

$$\text{máx}\{|\nu_\alpha(t)| : \alpha \in \Gamma_+(t)\} < \text{mín}\{\mu_\alpha(t) : \alpha \in \Gamma_+(t)\} \quad (\forall t \in T),$$

donde $\Gamma_+(t) = \{\alpha \in I_a(u_0(t)) \mid \mu_\alpha(t) > 0\}$. De esta forma se implicaría que

$$(\hat{p}, \hat{\mu}) := (p + q, \mu + \nu) \in \Lambda(L, x_0, u_0).$$

Los siguientes ejemplos ilustran cómo esta implicación, efectivamente, puede no ocurrir.

Ejemplo 3.4.1. Considere las siguientes funciones:

$$f(t,x,u) = u, \quad L(t,x,u) = tu(2-u)/2, \quad \varphi(u) = u^2 - 1.$$

Sea $x_0 \in X$ arbitrario y defina

$$u_0(t) := \begin{cases} 1 & \text{si} \quad t \in [-1,0] \\ -1 & \text{si} \quad t \in (0,1]\,. \end{cases}$$

Sea $(x_0, u_0, p, \mu) \in \mathcal{E}$. Por la Definición 3.1.3, tenemos que

$$\mu(t) \geq 0, \quad \dot{p}(t) = 0, \quad p(t) = t(1 - u_0(t)) + 2u_0(t)\mu(t) \quad (t \in [-1,1]).$$

Así, p es una constante que satisface

$$p = \begin{cases} 2\mu(t) & \text{si} \quad t \in [-1,0] \\ 2t - 2\mu(t) & \text{si} \quad t \in (0,1]\,. \end{cases}$$

Como $\mu(t) \geq 0$ para toda $t \in [-1,1]$, de la primera relación obtenemos $p \geq 0$ y de la segunda $p \leq 2t$ para todo $t \in (0,1]$, lo cual se cumple sólo si $p \leq 0$. Por lo tanto, $p \equiv 0$ y

$$\mu(t) = \begin{cases} 0 & \text{si} \quad t \in [-1,0] \\ t & \text{si} \quad t \in (0,1]\,. \end{cases}$$

Lo anterior implica que $\Lambda(L, x_0, u_0) = \{(p, \mu)\}$. Sin embargo,

$$(q(t), \nu(t)) = (2, u_0(t)) \not\equiv (0,0)$$

satisface las condiciones de 3.2.10, pues $\nu(t) \geq 0$ y $\nu(t)\varphi(u_0(t)) = 0$ para todo $t \in [-1,1]$ tal que $\mu(t) = 0$, i.e., para todo $t \in [-1,0]$. Por otro lado, $\dot{q}(t) = 0$ y $q(t) = 2u_0(t)\nu(t) = 2$ para todo $t \in [-1,1]$. Por lo tanto, (x_0, u_0) no es normal con respecto a $S_1(\mu)$.

Una situación similar ocurre en el siguiente ejemplo.

Ejemplo 3.4.2. Considere las siguientes funciones:

$$f(t,x,u) = u, \quad L(t,x,u) = b(t)u, \quad \varphi(u) = (u^2 - 1)/2$$

donde

$$b(t) := \begin{cases} 0 & \text{si} \quad t \in [-1, 0] \\ t^2 & \text{si} \quad t \in (0, 1]\,. \end{cases}$$

Sea $x_0 \in X$ arbitrario y defina

$$u_0(t) := \begin{cases} 1 & \text{si} \quad t \in [-1, 0] \\ -1 & \text{si} \quad t \in (0, 1]\,. \end{cases}$$

Notemos que, si $(p, \mu) \in \Lambda(L, x_0, u_0)$, entonces

$$\mu(t) \geq 0, \quad \dot{p}(t) = 0, \quad p(t) = b(t) + u_0(t)\mu(t) \quad (t \in [-1, 1]).$$

Por lo tanto, p es una constante que satisface

$$p = \begin{cases} \mu(t) & \text{si} \quad t \in [-1, 0] \\ t^2 - \mu(t) & \text{si} \quad t \in (0, 1]\,. \end{cases}$$

Como $\mu(t) \geq 0$ para todo $t \in [-1, 1]$, de la primera relación obtenemos que $p \geq 0$ y, de la segunda, $p \leq t^2$ para todo $t \in (0, 1]$ y entonces $p \leq 0$. Así $p \equiv 0$ y por tanto $\mu \equiv b$. Esto implica que (p, μ) es el único par tal que $(x_0, u_0, p, \mu) \in \mathcal{E}$.

Consideremos ahora al par (q, ν) con $q \equiv 1$ y $\nu(t) := u_0(t)$ $(t \in [-1, 1])$. Este satisface 3.2.10(i) pues

$$\nu(t) \geq 0 \text{ y } \nu(t)\varphi(u_0(t)) = 0$$

para todo $t \in [-1, 1]$ tal que $\mu(t) = 0$, esto es, todo $t \in [-1, 0]$. Más aún, $\dot{q}(t) = 0$ y $q(t) = u_0(t)\nu(t) = 1$ para todo $t \in [-1, 1]$ y entonces 3.2.10(ii) y 3.2.10(iii) se satisfacen. Como $(q, \nu) \not\equiv (0, 0)$, concluimos que (x_0, u_0) no es normal con respecto a $S_1(\mu)$.

Notemos que, en los ejemplos anteriores, la dinámica está dada por $f(t, x, u) = u$, lo que muestra que aún en el caso (más simple) de cálculo de variaciones, la normalidad con respecto a $S_1(\mu)$ no es una caracterización de la unicidad de los multiplicadores si las restricciones son puntuales. Dicho de otra manera, el resultado de Kyparisis no es válido para estos problemas.

Lo contrario ocurre con el resultado principal de Wachsmuth en [44], el cual sí tiene contraparte en este contexto. Denotemos por $\mathcal{F}(x_0, u_0)$ al conjunto de todas las L que satisfacen las hipótesis de suavidad dadas al inicio de este capítulo y tales que (x_0, u_0) resuelve el problema P(L) de minimizar

$$I(x,u) = \int_{t_0}^{t_1} L(t, x(t), u(t))dt$$

sujeto a que $(x, u) \in S$.

Teorema 3.4.3. *Sea $(x_0, u_0) \in S$. Entonces son equivalentes:*

a. (x_0, u_0) *es fuertemente normal.*

b. $\Lambda(L, x_0, u_0)$ *es unitario para toda $L \in \mathcal{F}(x_0, u_0)$.*

Demostración.

(a) $\Rightarrow$ (b): Es precisamente la Nota 3.4.1.

(b) $\Rightarrow$ (a): Para todo $\gamma \in R \cup Q$ y $t \in T$, sea $\mu_\gamma(t) = 1$ si $\gamma \in I_a(u_0(t))$ y $\mu_\gamma(t) = 0$ en caso contrario. Definamos $L(t, x, u) := -\langle \mu(t), \varphi(u) \rangle$. Notemos que, para todo $(x, u) \in S$, se cumple

$$I(x_0, u_0) = \int_{t_0}^{t_1} L(t, x_0(t), u_0(t))dt = 0 \leq \int_{t_0}^{t_1} L(t, x(t), u(t))dt = I(x, u)$$

y, por tanto, $L \in \mathcal{F}(x_0, u_0)$. Claramente, $(x_0, u_0, 0, \mu) \in \mathcal{E}$, i.e., $(0, \mu) \in \Lambda(L, x_0, u_0)$.

Observemos que, para este caso,

$$H(t, x, u, 0, \mu(t), 1) = -L(t, x, u) - \langle \mu(t), \varphi(u) \rangle = 0.$$

Sea $(q, \nu) \in X \times \mathcal{U}_q$ satisfaciendo

i. $\nu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ \forall t \in T)$;

ii. $\dot{q}(t) = -A^*(t)q(t)$ $(\forall t \in T)$;

iii. $B^*(t)q(t) = \varphi'^*(u_0(t))\nu(t)$ $(\forall t \in T)$.

Para probar que $(q, \nu) \equiv (0, 0)$ y poder concluir (a), definamos

$$\hat{\mu}_\gamma(t) := \mu_\gamma(t) + \frac{\nu_\gamma(t)}{\beta} \quad \text{y} \quad \hat{p}(t) := \frac{q(t)}{\beta} \quad (\gamma \in R \cup Q,\ t \in T)$$

donde

$$\beta = 1 + \max \{ |\nu_\alpha(t)| : \alpha \in I_a(u_0(t)),\ \forall t \in T \}.$$

Veamos que $(x_0, u_0, \hat{p}, \hat{\mu}) \in \mathcal{E}$. En efecto, 3.1.3(a) se cumple pues si, $\varphi_\alpha(u_0(t)) = 0$, entonces

$$\hat{\mu}_\alpha(t) = \frac{\beta + \nu_\alpha(t)}{\beta} \geq 0$$

y, si $\varphi_\alpha(u_0(t)) < 0$, entonces $\hat{\mu}_\alpha(t)\varphi_\alpha(u_0(t)) = 0$. Por 3.1.3 (b) y (c) se tiene que

$$\dot{\hat{p}}(t) = \frac{\dot{q}(t)}{\beta} = -f_x^*(\tilde{x}_0(t))\frac{q(t)}{\beta} = -f_x^*(\tilde{x}_0(t))\hat{p}(t) = -A^*(t)\hat{p}(t).$$

Similarmente,

$$\begin{aligned} B^*(t)\hat{p}(t) &= f_u^*(\tilde{x}_0(t))\hat{p}(t) = f_u^*(\tilde{x}_0(t))\frac{q(t)}{\beta} = \varphi'^*(\tilde{x}_0(t))\frac{\nu(t)}{\beta} \\ &= L_u^*(\tilde{x}_0(t)) + \varphi'^*(\tilde{x}_0(t))\hat{\mu}(t). \end{aligned}$$

De lo anterior concluimos que $(\hat{p}, \hat{\mu}) \in \Lambda(L, x_0, u_0)$. Así, por (b), $(\hat{p}, \hat{\mu}) \equiv (0, \mu)$ y en consecuencia $(q, \nu) \equiv (0, 0)$. □

3.4.2. Caracterización de la unicidad de los multiplicadores de Lagrange

Como ilustran los ejemplos 3.4.1 y 3.4.2, dado un multiplicador (p, μ), en contraste con los casos anteriores, la normalidad con respecto a $S_1(\mu)$ no es una caracterización de la unicidad de dicho multiplicador. Por ello, en esta sección introducimos una nueva condición que depende también de un multiplicador (p, μ), resulta ser una consecuencia de la normalidad con respecto a $S_1(\mu)$, implica la normalidad con respecto a S y es, en efecto, necesaria y suficiente para la unicidad del multiplicador de Lagrange.

Esta nueva caracterización constituye una importante contribución a la teoría de unicidad de los multiplicadores de Lagrange en el contexto de control óptimo.

Definición 3.4.4. *Dada $\mu \in \mathcal{U}_q$, decimos que un proceso $(x_0, u_0) \in S$ satisface la condición $\mathcal{K}(\mu)$ si, dado $(q, \nu) \in X \times \mathcal{U}_q$ satisfaciendo*

i. $\nu_\alpha(t) + \mu_\alpha(t) \geq 0$ *y* $\nu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ \forall t \in T)$*;*

ii. $\dot{q}(t) = -A^*(t)q(t)$ $(\forall t \in T)$*;*

iii. $B^*(t)q(t) = \varphi'^*(u_0(t))\nu(t)$ $(\forall t \in T)$*,*

se tiene que $q \equiv 0$ lo que, a su vez, implica que $\nu \equiv 0$.

Nuestro siguiente resultado muestra que, en efecto, esta condición caracteriza la unicidad de un multiplicador de Lagrange $(p, \mu) \in \Lambda(L, x_0, u_0)$.

Teorema 3.4.5. *Sea $(x_0, u_0) \in S$ para el cual existe $(p, \mu) \in \Lambda(L, x_0, u_0)$. Entonces* *los siguientes enunciados son equivalentes:*

a. (x_0, u_0) *satisface la condición* $\mathcal{K}(\mu)$.

b. $\Lambda(L, x_0, u_0) = \{(p, \mu)\}$.

Demostración.

(a) $\Rightarrow$ (b): Supongamos que $(\bar{p}, \bar{\mu}) \in \Lambda(L, x_0, u_0)$ y sea $(q, \nu) := (\bar{p} - p, \bar{\mu} - \mu)$. Mostraremos que (q, ν) satisface las condiciones de la Definición 3.4.4. Si $\alpha \in R$ y $t \in T$, entonces

$$\nu_\alpha(t) + \mu_\alpha(t) = \bar{\mu}_\alpha(t) \geq 0 \quad \text{y} \quad \nu_\alpha(t)\varphi_\alpha(u_0(t)) = [\bar{\mu}_\alpha(t) - \mu_\alpha(t)]\, \varphi_\alpha(u_0(t)) = 0$$

y, por tanto, 3.4.4(i) se satisface. Las condiciones 3.4.4(ii) y (iii) se siguen de

$$\dot{q}(t) = \dot{\bar{p}}(t) - \dot{p}(t) = -A^*(t)q(t) \quad (\forall t \in T),$$

$$B^*(t)q(t) = B^*(t)(\bar{p}(t) - p(t)) = \varphi'^*(u_0(t))\nu(t) \quad (\forall t \in T).$$

Como (x_0, u_0) satisface la condición $\mathcal{K}(\mu)$, $(q, \nu) \equiv (0, 0)$.

(b) $\Rightarrow$ (a): $\neg$(a) $\Rightarrow$ $\neg$(b): Supongamos que (x_0, u_0) no satisface la condición $\mathcal{K}(\mu)$. Entonces existe $(q, \nu) \in X \times \mathcal{U}_q$ satisfaciendo 3.4.4(i)-(iii). Como se verifica fácilmente, $(\hat{p}, \hat{\mu}) := (p + q, \mu + \nu) \in \Lambda(L, x_0, u_0)$. En efecto,

$$\hat{\mu}_\alpha(t) \geq 0 \quad \text{y} \quad \hat{\mu}_\alpha(t)\varphi_\alpha(u_0(t)) = 0 \quad (\alpha \in R,\ t \in T)$$

y, para todo $t \in T$,

$$\dot{\hat{p}}(t) = \dot{p}(t) + \dot{q}(t) = -A^*(t)\hat{p}(t) + L_x^*(\tilde{x}_0(t)),$$

$$B^*(t)\hat{p}(t) = B^*(t)(p(t) + q(t)) = L_u^*(\tilde{x}_0(t)) + \varphi'^*(u_0(t))\hat{\mu}(t).$$

Por lo tanto, como $(\hat{p}, \hat{\mu}) \not\equiv (p, \mu)$, $\Lambda(L, x_0, u_0) \neq \{(p, \mu)\}$. □

La siguiente proposición muestra la conexión entre $\mathcal{K}(\mu)$ y las nociones de normalidad débil y normalidad con respecto a $S_1(\mu)$.

Proposición 3.4.6. *Sea* $(x_0, u_0) \in S$ *y supongamos que* $(p, \mu) \in \Lambda(L, x_0, u_0)$. *Entonces,* (x_0, u_0) *es normal con respecto a* $S_1(\mu)$ $\Rightarrow$ (x_0, u_0) *satisface la condición* $\mathcal{K}(\mu)$ $\Rightarrow$ (x_0, u_0) *es normal con respecto a* S.

Demostración. Supongamos que (x_0, u_0) es normal con respecto a $S_1(\mu)$ y (q, ν)

satisface las condiciones 3.4.4(i)-(iii). La primera implicación se seguirá si mostramos que (q, ν) satisface 3.2.10(i)-(iii). Así, es suficiente probar que 3.4.4(i) implica 3.2.10(i), lo cual se cumple porque $\nu_\alpha(t) + \mu_\alpha(t) \geq 0$ para todo $\alpha \in R$ y para todo $t \in T$, implicando así que $\nu_\alpha(t) \geq 0$ para todo $\alpha \in R$ para el cual $\mu_\alpha(t) = 0$ $(\forall t \in T)$.

Para la segunda implicación, si (x_0, u_0) satisface $\mathcal{K}(\mu)$ y (q, ν) satisface 3.1.2(i)-(iii), como antes, es suficiente probar que 3.1.2(i) implica 3.4.4(i), lo cual es claro porque $\mu_\alpha(t) \geq 0$ y $\nu_\alpha(t) \geq 0$ $(\alpha \in R,\ t \in T)$. □

Nuestro siguiente resultado muestra qué hipótesis adicional se requiere para que la condición $\mathcal{K}(\mu)$ y la normalidad con respecto a $S_1(\mu)$ sean equivalentes. En otras palabras, da una condición suficiente para que la caracterización de Kyparisis sea válida para (P).

Nota 3.4.7. *Sea $(x_0, u_0) \in S$ y supongamos que $(p, \mu) \in \Lambda(L, x_0, u_0)$. Como en el Teorema 3.4.2, definimos*

$$\Gamma_+(t) = \{\alpha \in I_a(u_0(t)) \mid \mu_\alpha(t) > 0\} \quad (t \in T).$$

Si $\inf_{t \in T} \{\mu_\alpha(t) \mid \alpha \in \Gamma_+(t)\} > 0$, entonces (x_0, u_0) es normal con respecto a $S_1(\mu) \Leftrightarrow$ (x_0, u_0) satisface $\mathcal{K}(\mu)$.

Demostración. La primera implicación la tenemos por la Proposición 3.4.6. Para el recíproco observe que, si (x_0, u_0) no es normal con respecto a $S_1(\mu)$, entonces existe (q, ν) satisfaciendo las condiciones 3.2.10(i)-(iii) con $q \not\equiv 0$ y, en vista de nuestra hipótesis, ν puede ser elegido de forma que

$$\max\{|\nu_\alpha(t)| : \alpha \in \Gamma_+(t)\} < \min\{\mu_\alpha(t) : \alpha \in \Gamma_+(t)\} \quad (\forall t \in T).$$

Sea $(\hat{p}, \hat{\mu}) := (p + q, \mu + \nu)$. Probaremos que $(\hat{p}, \hat{\mu}) \in \Lambda(L, x_0, u_0)$ y, como $(\hat{p}, \hat{\mu}) \not\equiv (p, \mu)$, (p, μ) no es único, lo que por Teorema 3.4.5 implica que (x_0, u_0) no satisface $\mathcal{K}(\mu)$.

Notemos primero que, para todo $t \in T$, se tiene que

$$\dot{\hat{p}}(t) = \dot{p}(t) + \dot{q}(t) = -A^*(t)\hat{p}(t) + L_x^*(\tilde{x}_0(t)),$$

$$B^*(t)\hat{p}(t) = B^*(t)(p(t) + q(t)) = L_u^*(\tilde{x}_0(t)) + \varphi'^*(u_0(t))\hat{\mu}(t)$$

y, por tanto, 3.2.10(ii) y (iii) se satisfacen. Para ver 3.2.10(i), consideremos $\alpha \in R$ y

$t \in T$. Si $\mu_\alpha(t) = 0$, entonces $\hat{\mu}_\alpha(t) = \nu_\alpha(t) \geq 0$. Si $\mu_\alpha(t) > 0$, entonces

$$\hat{\mu}_\alpha(t) = \mu_\alpha(t) + \nu_\alpha(t) \geq \min_{i \in \Gamma_+(t)} \mu_i(t) + \nu_\alpha(t) > \max_{i \in \Gamma_+(t)} |\nu_i(t)| + \nu_\alpha(t) \geq 0.$$

Finalmente notemos que, si $\varphi_\alpha(u_0(t)) = 0$, entonces $\hat{\mu}_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ y, si $\varphi_\alpha(u_0(t)) < 0$, por 3.2.10(i) $\mu_\alpha(t) = 0$ y por 3.4.4(i),

$$0 = \nu_\alpha(t)\varphi_\alpha(u_0(t)) = \hat{\mu}_\alpha(t)\varphi_\alpha(u_0(t)). \qquad \square$$

3.4.3. Unicidad y τ-regularidad

Como con las nociones de normalidad, buscamos caracterizar la condición $\mathcal{K}(\mu)$ en términos de los conos τ introducidos en [39] ya que, en sus definiciones originales, éstas dependen de un sistema definido por parejas $(q, \nu) \in X \times \mathcal{U}_q$, mientras que las caracterizaciones dependen explícitamente sólo de una de ellas lo que, generalmente, permite que se verifiquen más fácilmente.

Con este fin, introducimos un nuevo tipo de regularidad de la siguiente manera.

Definición 3.4.8. *Dados $(x_0, u_0) \in S$ y $\mu \in \mathcal{U}_q$, diremos que (x_0, u_0) es $\tau(\mu)$-regular si $z \equiv 0$ es la única solución en X del sistema*

$$\dot{z}(t) = -A^*(t)z(t), \quad [z^*(t)B(t) + \mu^*(t)\varphi'(u_0(t))]h \leq 0 \ \forall h \in \tau(u_0(t)) \quad (\forall t \in T).$$

Teorema 3.4.9. *Sean $(x_0, u_0) \in S$ y $\mu \in \mathcal{U}_q$ con $\mu_\alpha(t) \geq 0$ y $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ \forall t \in T)$. Entonces las siguientes afirmaciones son equivalentes:*

a. (x_0, u_0) *es $\tau(\mu)$-regular.*

b. (x_0, u_0) *satisface la condición $\mathcal{K}(\mu)$.*

Demostración. Sea $\sigma(t) := \mu^*(t)\varphi'(u_0(t))$ $(\forall t \in T)$.

(a) $\Rightarrow$ (b): Supongamos que $(q, \nu) \in X \times \mathcal{U}_q$ satisface 3.4.4(i)-(iii). Si $\lambda := \nu + \mu$, notemos que $\lambda_\alpha(t) \geq 0$ y $\lambda_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ \forall t \in T)$. Además

$$\dot{q}(t) = -A^*(t)q(t), \quad B^*(t)q(t) = \varphi'^*(u_0(t))(\lambda(t) - \mu(t)) \quad (\forall t \in T).$$

Sea $h \in \tau(u_0(t))$. Tenemos entonces que

$$q^*(t)B(t)h = (\lambda^*(t) - \mu^*(t))\varphi'(u_0(t))h \quad (\forall t \in T)$$

y, por tanto,

$$[q^*(t)B(t)+\sigma(t)]\,h=\sum_{i=1}^{q}\lambda_i(t)\varphi_i'(u_0(t))h\leq 0$$

implicándose por (a) que $q\equiv 0$.

(b) $\Rightarrow$ (a): Sea $z\in X$ tal que

$$\dot{z}(t)=-A^*(t)z(t),\quad [z^*(t)B(t)+\sigma(t)]h\leq 0 \text{ para todo } h\in\tau(u_0(t))\quad (t\in T).$$

Como en la prueba de 3.3.2, para cada $t\in T$, definimos

$$\hat{\varphi}:=(\varphi_{i_1},\ldots,\varphi_{i_p})^*\quad \text{donde } I_a(u_0(t))\cup Q=\{i_1,\ldots,i_p\}$$

y $\hat{\nu}:=(\nu_{i_1},\ldots,\nu_{i_p})^*$ por

$$\hat{\nu}(t):=\phi^{-1}(t)\hat{\varphi}'(u_0(t))B^*(t)z(t)\ (\forall t\in T)$$

donde $\phi(t)=\hat{\varphi}'(u_0(t))\hat{\varphi}'^*(u_0(t))$ $(\forall t\in T)$. Notemos que, dado que $\phi(t)$ es simétrica, también ϕ^{-1} lo es. Sea $\nu(t)=(\nu_1(t),\ldots,\nu_q(t))^*$ donde

$$\nu_\alpha(t):=\begin{cases}\hat{\nu}_{i_r}(t) & \alpha=i_r,\ r=1,\ldots,p\\ 0 & \text{en otro caso.}\end{cases}$$

Claramente, $\nu_\alpha(t)\varphi_\alpha(u_0(t))=0$ $(\alpha\in R,\ t\in T)$ y

$$\hat{\varphi}'^*(u_0(t))\hat{\nu}(t)=\begin{pmatrix}\sum_{j=1}^{p}\dfrac{\partial\varphi_{i_j}}{\partial u_1}(u_0(t))\hat{\nu}_{i_j}(t)\\ \vdots\\ \sum_{j=1}^{p}\dfrac{\partial\varphi_{i_j}}{\partial u_m}(u_0(t))\hat{\nu}_{i_j}(t)\end{pmatrix}=\varphi'^*(u_0(t))\nu(t)$$

y transponiendo obtenemos que, para $t\in T$, $\hat{\nu}^*(t)\hat{\varphi}'(u_0(t))=\nu^*(t)\varphi'(u_0(t))$. Luego,

$$G(t):=I_{m\times m}-\hat{\varphi}'^*(u_0(t))\phi^{-1}(t)\hat{\varphi}'(u_0(t))$$

satisface que $\hat{\varphi}'(u_0(t))G(t)=0$ $(\forall t\in T)$. Si $h_k(t)$ $(k=1,\ldots,m)$ denota la k-ésima columna de $G(t)$, tendremos que

$$\varphi_{i_j}'(u_0(t))h_k(t)=0\quad (j=1,\ldots,p,\ k=1,\ldots,m).$$

Así, $h_k(t) \in \tau(u_0(t))$ y, por tanto, $[z^*(t)B(t) + \sigma(t)]\, h_k(t) \leq 0$ $(k = 1, \ldots, m)$, pero $-h_k(t)$ también satisface esta desigualdad, por lo que $[z^*(t)B(t) + \sigma(t)]\, h_k(t) = 0$. Notemos además que $\sigma(t)G(t) = 0$, de donde obtenemos

$$0 = [z^*(t)B(t) + \sigma(t)]\, G(t) = z^*(t)B(t) - \nu^*(t)\varphi'(u_0(t)).$$

Por lo tanto,

$$B^*(t)z(t) = \varphi'^*(u_0(t))\nu(t) \quad (\forall t \in T).$$

Por el Lema 3.3.1 concluimos que ν es continua a trozos y, por tanto, también $\nu + \mu$. Resta verificar que $\nu_\alpha(t) + \mu_\alpha(t) \geq 0$ $(\alpha \in R,\ \forall t \in T)$. Para ello, consideremos la matriz $\hat{C}(t)$ de la Proposición 3.3.3, definida por $\hat{C}(t) := \phi^{-1}(t)\hat{\varphi}'(u_0(t))$. Tenemos entonces que

$$\hat{C}(t)\hat{\varphi}'^*(u_0(t)) = I_{p\times p} = \hat{\varphi}'(u_0(t))\hat{C}^*(t).$$

Por lo tanto, si $\hat{c}_j(t)$ denota la j-ésima columna de $\hat{C}^*(t)$ y $\{e_j\}$ la base canónica en $\mathbb{R}^p$ $(j = 1, \ldots, p)$, entonces

$$(\varphi'_{i_j}(u_0(t))\hat{c}_1(t), \ldots, \varphi'_{i_j}(u_0(t))\hat{c}_p(t)) = e_j^* \quad (j = 1, \ldots, p).$$

Como hemos visto, si $i_j \in I_a(u_0(t))$, entonces

$$\varphi'_k(u_0(t))(-\hat{c}_j(t)) = \begin{cases} -1 & \text{si} \quad k = i_j \\ 0 & \text{si} \quad k \neq i_j \end{cases}$$

implicando que $-\hat{c}_j(t) \in \tau(u_0(t))$. Por tanto $[z^*(t)B(t) + \sigma(t)]\, \hat{c}_j(t) \geq 0$ $(\forall t \in T)$, pero

$$\hat{\nu}^*(t) = z^*(t)B(t)\hat{C}^*(t) \text{ y } \hat{\mu}^*(t) = \sigma(t)\hat{C}^*(t)$$

donde $\hat{\mu}(t) := (\mu_{i_1}, \ldots, \mu_{i_p})^*$. Por lo tanto,

$$\hat{\nu}_\alpha(t) + \hat{\mu}_\alpha(t) \geq 0 \quad \text{para todo } \alpha \in I_a(u_0(t))$$

y esto completa la demostración. □

3.5. Sobre condiciones de segundo orden

En la sección 3.2 hemos explicado por qué sería conveniente establecer condiciones de segundo orden para (P) en términos de los conos determinados por las restricciones tangenciales ($\mathcal{R}_S$, $\mathcal{R}_{S_1}$ o $\mathcal{R}_{S_0}$) y no de los conjuntos de tangentes secuenciales y curvilíneos ($\mathcal{T}_{S_1}$ y C_{S_1}) como se hizo anteriormente. Sin embargo, bajo hipótesis de normalidad, no es mucho lo que podemos decir sobre el comportamiento de $J((x_0, u_0, p, \mu); (\cdot, \cdot))$ en ninguno de estos conos.

Uno de los resultados que relacionan estos conceptos (normalidad, condiciones en $\mathcal{R}_S$) es el siguiente, el cual se obtuvo en [19, Teorema 6.7] reduciendo el problema original a uno que involucra únicamente restricciones de igualdad.

Teorema 3.5.1. *Sea $(x_0, u_0) \in S$ y suponga que existe $(p, \mu) \in X \times \mathcal{U}_q$ tal que $(x_0, u_0, p, \mu) \in \mathcal{E}$. Si (x_0, u_0) resuelve (P) y es un proceso normal con respecto a S_0, entonces*

$$J((x_0, u_0, p, \mu); (y, v)) \geq 0 \text{ para todo } (y, v) \in \mathcal{R}_{S_0}(x_0, u_0)$$

donde, recordemos,

$$\mathcal{R}_{S_0}(x_0, u_0) := \{(y, v) \in \mathcal{L}(x_0, u_0) \mid v(t) \in \tau_0(u_0(t)) \ (\forall t \in T)\}.$$

Este resultado es aplicable a una clase muy particular de problemas pues, como se muestra en el siguiente ejemplo, la conclusión de este teorema puede no ser válida si debilitamos la hipótesis y suponemos normalidad débil y no normalidad fuerte.

Ejemplo 3.5.1. *Sean $a > 0$, $k > 1$. Consideremos el problema de minimizar*

$$I(x, u) = \int_0^1 \{u_2(t) - u_1(t)\} dt$$

sujeto a $(x, u) \in Z$ y

$$\begin{cases} \dot{x}(t) = 2u_2(t) - (k+1)u_1(t) + au_3^2(t) \quad (t \in [0, 1]); \\ x(0) = x(1) = 0; \\ u_2(t) \geq u_1(t), \ ku_1(t) \geq u_2(t) \quad (t \in [0, 1]). \end{cases}$$

En este caso, tenemos

$$H(t,x,u,p,\mu) =$$
$$p(2u_2 - (k+1)u_1 + au_3^2) + u_1 - u_2 - \mu_1(u_1 - u_2) - \mu_2(u_2 - ku_1),$$

por lo que $H_x(t,x,u,p,\mu) \equiv 0$,

$$H_u(t,x,u,p,\mu) = (-p(k+1) - \mu_1 + k\mu_2 + 1, 2p + \mu_1 - \mu_2 - 1, 2apu_3),$$

$$H_{uu}(t,x,u,p,\mu) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2ap \end{pmatrix}$$

por lo cual, para todo $(x,u,p,\mu) \in Z \times X \times \mathcal{U}_2$ y $(y,v) \in Z$,

$$J((x,u,p,\mu);(y,v)) = -\int_0^1 2ap(t)v_3^2(t)dt.$$

Es fácil verificar que $(x_0,u_0) \equiv (0,0)$ es una solución de (P). Probaremos que si $\mu^* = (\mu_1,\mu_2) \equiv (0,1)$ y $p \equiv 1$, entonces $(x_0,u_0,p,\mu) \in \mathcal{E}$. En efecto, 3.1.3 (a) se satisface de manera trivial pues, $I_a(u_0(t)) \equiv \{1,2\}$. Como $A(t) = f_x(\tilde{x}_0(t)) = 0$ y $L_x(\tilde{x}_0(t)) \equiv 0$, 3.1.3 (b) implica que $\dot{p}(t) \equiv 0$, lo cual se cumple puesto que p es constante. Finalmente, como $\varphi_1'(u) = (1,-1,0)$, $\varphi_2'(u) = (-k,1,0)$ y

$$B(t) = f_u(\tilde{x}_0(t)) = (-(k+1), 2, 0),$$

tenemos que

$$\begin{aligned} p^*(t)B(t) &= (-(k+1),2,0) = (-1,1,0) + (0,1)\begin{pmatrix} 1 & -1 & 0 \\ -k & 1 & 0 \end{pmatrix} \\ &= L_u(\tilde{x}_0(t)) + \mu^*(t)\varphi'(u_0(t)), \end{aligned}$$

y, transponiendo, obtenemos 3.1.3(c). Por otra parte,

$$\tau_0(u_0(t)) = \{h \in \mathbb{R}^3 \mid h_1 - h_2 = 0,\ h_2 - kh_1 = 0\},$$

$$\tau_1(u_0(t),\mu(t)) = \{h \in \mathbb{R}^3 \mid h_1 - h_2 \le 0,\ h_2 - kh_1 = 0\},$$

$$\tau(u_0(t)) = \{h \in \mathbb{R}^3 \mid h_1 - h_2 \le 0,\ h_2 - kh_1 \le 0\}.$$

Note que el sistema

$$\dot{z}(t) = -A^*(t)z(t) = 0, \quad z^*(t)B(t)h = z(t)(-(k+1)h_1 + 2h_2) = 0$$

para todo $(h_1, h_2, h_3) \in \tau_0(u_0(t))$ $(\forall t \in T)$ tiene soluciones no nulas. Por lo tanto, (x_0, u_0) no es normal con respecto a S_0.

Si consideramos $\tau_1(u_0(t), \mu(t))$, el sistema anterior se transforma en

$$\dot{z}(t) = 0, \quad z(t)(h_2 - h_1) \leq 0$$

para todo $(h_1, h_2, h_3) \in \tau_1(u_0(t), \mu(t))$ $(\forall t \in T)$, el cual también tiene soluciones no nulas por lo que (x_0, u_0) no es normal con respecto a $S_1(\mu)$.

Finalmente, el sistema

$$\dot{z}(t) = 0, \quad z(t)(-(k+1)h_1 + 2h_2) \leq 0$$

para todo $(h_1, h_2, h_3) \in \tau(u_0(t))$ $(\forall t \in T)$, tiene como única solución $z \equiv 0$ pues, para cada $t \in T$, $(1, 1, 0)$ y $(1, k, 0)$ pertenecen a $\tau(u_0(t))$ lo que implica $z(t) \geq 0$ y $z(t) \leq 0$. Por lo tanto, (x_0, u_0) es normal con respecto a S.

Sean $v \equiv (0, 0, 1)$, $y \equiv 0$. Notemos que $(y, v) \in \mathcal{R}_{S_0}$, pues satisface

$$y(0) = y(1) = 0, \;\; \dot{y}(t) = -(k+1)v_1(t) + 2v_2(t), \;\; v(t) \in \tau_0(u_0(t)) \quad (\forall t \in T)$$

y $J((x_0, u_0, p, \mu); (y, v)) = -2a < 0$.

Notemos que en este ejemplo $\Gamma_+(t) \equiv \{2\}$, por lo que $\inf_{t \in T}\{\mu_2(t)\} = 1 > 0$. Por la Nota 3.4.7 podemos concluir que el multiplicador $(1, (0, 1))$ no es único. De hecho, el sistema de la Definición 3.4.9 de $\tau(\mu)$-regularidad es

$$\dot{z}(t) = 0, \quad (-z(t)(k+1) - k)h_1 + (2z(t) + 1)h_2 \leq 0 \tag{3.5.1}$$

para todo $(h_1, h_2, h_3) \in \tau(u_0(t))$ $(\forall t \in T)$, tiene soluciones no nulas, y

$$\Lambda(u_2(t) - u_1(t), (0, 0)) = \{(q, \lambda) \mid q = 1 + z, (\lambda_1, \lambda_2) = (0, 1) + (\nu_1, \nu_2)\}$$

donde z es solución del sistema (3.5.1) y $\nu = \phi^{-1}\varphi' B^* z$. Note que en este caso particular las funciones son constantes.

Nuestro siguiente ejemplo ilustra la situación opuesta, una solución fuertemente normal para la cual existe $(y, v) \in \mathcal{R}_S(x_0, u_0)$ tal que $J((x_0, u_0, p, \mu); (y, v)) < 0$.

Ejemplo 3.5.2. *Consideremos el problema de minimizar*

$$I(x, u) = \int_0^1 -\exp(-u_2(t))dt$$

sujeto a $(x, u) \in Z$ *y*

$$\begin{cases} \dot{x}(t) = u_1(t) + u_2^2(t) \quad (t \in [0, 1]); \\ x(0) = x(1) = 0; \\ u_2(t) \geq 0 \quad (t \in [0, 1]). \end{cases}$$

En este caso, $H(t, x, u, p, \mu) = p(u_1 + u_2^2) + \mu_1 u_2 + e^{-u_2}$, y tenemos

$$H_u(t, x, u, p, \mu) = (p, 2pu_2 + \mu_1 - e^{-u_2}),$$

$$H_{uu}(t, x, u, p, \mu) = \begin{pmatrix} 0 & 0 \\ 0 & 2p + e^{-u_2} \end{pmatrix}$$

de modo que, para todo $(x, u, p, \mu) \in Z \times \mathcal{U}_1$ y $(y, v) \in Z$,

$$J((x, u, p, \mu); (y, v)) = -\int_0^1 (2p(t) + e^{-u_2(t)})v_2^2(t)dt.$$

Claramente, $(x_0, u_0) \equiv (0, 0)$ resuelve el problema. Como $f_u(t, x, u) = (1, 2u_2)$, $L_u(t, x, u) = (0, e^{-u_2})$ y $\varphi'(u) = (0, -1)$, (x_0, u_0, p, μ) es un extremo si $\dot{p}(t) = 0$ y

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} p(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ -\mu(t) \end{pmatrix}.$$

Así, (x_0, u_0, p, μ) es un extremo si $(p, \mu) \equiv (0, 1)$. Observemos que

$$\tau_0(u_0(t)) = \tau_1(u_0(t), \mu(t)) = \{h \in \mathbb{R}^2 \mid h_2 = 0\},$$

$$\tau(u_0(t)) = \{h \in \mathbb{R}^2 \mid -h_2 \leq 0\}.$$

Entonces, $z \equiv 0$ es la única solución del sistema

$$\dot{z}(t) = 0, \quad z(t)h_1 = 0 \quad \text{para todo } (h_1, h_2) \in \tau_0(u_0(t)) \quad (\forall t \in T)$$

y, por lo tanto, (x_0, u_0) es normal con respecto a S_0. Luego, si $(y, v) \equiv (0, (0, 1))$, se tiene que $(y, v) \in \mathcal{R}_S(x_0, u_0)$ pues $v(t) \in \tau(u_0(t))$ y $\dot{y}(t) = v_1(t)$ $(\forall t \in T)$, $y(0) = y(1) = 0$, pero $J((x_0, u_0, p, \mu); (y, v)) = -1 < 0$.

En vista de estos ejemplos, el resultado del Teorema 3.5.1 es muy débil en comparación con el resultado principal que se tiene en el caso en dimensión finita (Teorema 1.2.12), generalizado al caso de cálculo de variaciones con restricciones isoperimétricas (Teoremas 2.1.6), en los cuales se supone normalidad con respecto a S_1 y se puede asegurar que la segunda variación es no negativa en el conjunto de restricciones tangenciales correspondiente a S_1 mientras que, en 3.5.1, se supone normalidad con respecto a S_0 (que en general es una hipótesis más fuerte que la normalidad con respecto a S_1) y la condición de la segunda variación se satisface en $\mathcal{R}_{S_0}$ (un conjunto, en general, menor que $\mathcal{R}_{S_1}$).

3.5.1. Una conjetura sobre S_1-normalidad

Con base en la observación anterior y el Teorema 3.4.2, es natural hacer la siguiente conjetura (la contraparte del Teorema 1.2.12), planteada por primera vez en [31].

Teorema 3.5.2. *Sea $(x_0, u_0) \in S$ para el cual existe $(p, \mu) \in X \times \mathcal{U}_q$ tal que $(x_0, u_0, p, \mu) \in \mathcal{E}$. Si (x_0, u_0) resuelve (P) y es normal con respecto a $S_1(\mu)$, entonces*

$$J((x_0, u_0, p, \mu); (y, v)) \geq 0 \quad \textit{para todo } (y, v) \in \mathcal{R}_{S_1}(x_0, u_0).$$

Algunas pruebas parciales del Teorema 3.5.2 pueden encontrarse en [31, Teorema 3.3], donde se supone que el conjunto U es convexo, [46, Teorema 4.1], donde las funciones $I_a(u_0(t))$, $\Gamma_0(t)$ y $\Gamma_+(t)$ son constantes a trozos y, más recientemente, en [15, 16, 38], donde la función del costo no depende de x, es decir, $L(t, x, u) = L(t, u)$.

Recordemos que, en los casos previos, la implicación que permitió probar la condición de segundo orden fue la siguiente:

Normalidad (con respecto a S) implica regularidad (con respecto a S).

Notemos que si éste fuera el caso para (P), por el Teorema 3.2.4 o por el Teorema 3.2.5 (sustituyendo regularidad por $\mathcal{T}$-regularidad), obtendríamos la conclusión del Teorema 3.5.2. Sorprendentemente, nuestro último ejemplo prueba que, en este contexto, no necesariamente se cumple esta implicación.

3.5.2. Normalidad no implica regularidad

Ejemplo 3.5.3. *Considere el problema de minimizar* $\int_{-1}^{1} L(t,x(t),u(t))dt$ *sujeto a*

$$\begin{cases} x(-1) = 1/2, \ x(1) = 0 \\ \dot{x}(t) = u(t) \leq 0 \quad (t \in [-1,1]). \end{cases}$$

Para este caso, tenemos $T = [-1,1]$, $n = m = 1$,

$$D = \{(x,u) \in X \times \mathcal{U}_1 \mid \dot{x}(t) = u(t) \ (t \in T), \ x(-1) = 1/2, \ x(1) = 0\},$$

$$S = \{(x,u) \in D \mid u(t) \leq 0 \ (\forall t \in T)\},$$

$$U = \{u \in \mathbb{R} \mid \varphi(u) \leq 0\} \ \text{ con } \varphi(u) = u.$$

Considere los siguientes arcos

$$x_0(t) := \begin{cases} t^2/2 & \text{si} \quad t \in [-1,0) \\ 0 & \text{si} \quad t \in [0,1] \end{cases} \qquad y(t) := \begin{cases} t+1 & \text{si} \quad t \in [-1,0) \\ 1-t & \text{si} \quad t \in [0,1] \end{cases}$$

de modo que

$$\dot{x}_0(t) = u_0(t) = \begin{cases} t & \text{si} \quad t \in [-1,0) \\ 0 & \text{si} \quad t \in [0,1] \end{cases} \qquad \dot{y}(t) = v(t) = \begin{cases} 1 & \text{si} \quad t \in [-1,0) \\ -1 & \text{si} \quad t \in [0,1] \,. \end{cases}$$

Claramente, para cualquier función $L(t,x,u)$, (x_0,u_0) es un proceso admisible para el cual $\varphi(u_0(t))$ únicamente se anula en el subintervalo $[0,1]$.

Es claro también que

$$(y,v) \in \mathcal{R}_S(x_0,u_0) = \{(y,v) \in \mathcal{L}(x_0,u_0) \mid v(t) \leq 0 \ (\forall t \in [0,1])\},$$

donde, para este caso,

$$\mathcal{L}(x_0,u_0) = \{(y,v) \in X \times \mathcal{U}_q \mid \dot{y}(t) = v(t) \ (\forall t \in T), \ y(-1) = y(1) = 0\}.$$

De hecho, note que v satisface de manera estricta la desigualdad que define a $\mathcal{R}_S(x_0, u_0)$.

Por otra parte, $(p, \mu) \equiv (0, 0)$ es la única solución en $X \times \mathcal{U}_1$ del sistema

i. $\mu(t)u_0(t) = 0 \quad (\forall t \in T)$,

ii. $\dot{p}(t) = 0 \quad (\forall t \in T)$,

iii. $p(t) = \mu(t) \quad (\forall t \in T)$

pues, por (i), $\mu(t) = 0$ ($t \in [-1, 0)$). Por (ii), $p \equiv k$ una constante y, por (iii), $k = \mu(t)$ ($\forall t \in T$) y, consecuentemente, $k = 0$. Por lo tanto, (x_0, u_0) es fuertemente normal (en particular normal con respecto a S).

Mostraremos por contradicción que $(y, v) \notin \mathcal{T}_S(x_0, u_0)$. Supongamos lo contrario, y sean $\{z_k := (x_k, u_k)\} \subset S$ y $\{\epsilon_k > 0\}$ sucesiones tales que

$$\lim_{k\to\infty} \epsilon_k = 0, \quad \lim_{k\to\infty} \frac{z_k - z_0}{\epsilon_k} = \rho$$

donde $z_0 = (x_0, u_0)$, $\rho = (y, v)$ y la norma que consideramos en $X \times \mathcal{U}_1$ es la norma definida por la ecuación (2.2.1).

Consideremos las sucesiones

$$\rho_k(t) = (y_k(t), v_k(t)) := \frac{z_k(t) - z_0(t)}{\epsilon_k} - \rho(t), \quad w_k := \sup_{t\in[-1,0)} |v_k(t)|$$

y observemos que

$$\begin{aligned} \lim_{k\to\infty} w_k &\le \lim_{k\to\infty} \sup_{t\in T} |v_k(t)| \\ &\le \lim_{k\to\infty} [\sup_{t\in T} (|y_k(t)| + |\dot{y}_k(t)|) + \sup_{t\in T} |v_k(t)|] = \lim_{k\to\infty} |\rho_k(t)| = 0. \end{aligned} \tag{3.5.2}$$

Sean $0 < \eta < 1$ y $N \in \mathbb{N}$ tales que $k \ge N$ implique $|v_k(t)| \le w_k < \eta$ para todo $t \in [-1, 0)$. Tenemos entonces que

$$|v_k(t)| = \left| \frac{u_k(t) - t}{\epsilon_k} - 1 \right| = \left| \frac{u_k(t) - t - \epsilon_k}{\epsilon_k} \right| < \eta \quad (t \in [-1, 0))$$

lo que implica que $-\eta\epsilon_k < u_k(t) - t - \epsilon_k$ y, como $\eta < 1$, vemos que

$$u_k(t) > t + \epsilon_k(1 - \eta) > 0 \quad \text{para todo } t \in (\epsilon_k(\eta - 1), 0) \cap [-1, 0)\,,$$

contradiciendo así que $z_k \in S$. Por lo tanto $(y, v) \notin \mathcal{T}_S(x_0, u_0)$. $\square$

Como en los ejemplos 3.4.1 y 3.4.2, en el caso anterior la dinámica está dada por $f(t, x, u) = u$, lo que implica que $u_0 = \dot{x}$, $v = \dot{y}$, por lo que el problema anterior puede ser considerado dentro del contexto de cálculo de variaciones y, dado que al considerar la norma débil en X ($\|x\| = \sup_{t\in T}(x(t)^2 + \dot{x}(t)^2)^{1/2}$) la desigualdad (3.5.2) también se satisface, tenemos que $y \notin \mathcal{T}_S(x_0)$, lo que prueba que la implicación normal $\Rightarrow$ $\mathcal{T}$-regular (con respecto a S) no necesariamente es válida, aun en ese problema más sencillo.

De manera semejante podemos probar que, dada cualquier familia de funciones $(x(t, \epsilon), u(t, \epsilon)) \in X \times \mathcal{U}_1$, definida en $T \times [0, \delta)$ para algún $\delta > 0$, tal que

a. $(x(t, 0), u(t, 0)) = (x_0(t), u_0(t))$ $(\forall t \in T)$;

b. $(x_\epsilon(t, 0), u_\epsilon(t, 0)) = (y(t), v(t))$ $(\forall t \in T)$;

c. $x(t, \cdot)$ es de clase C^2; $u(t, \cdot)$ es de clase C^2 a trozos,

entonces, para todo $\epsilon > 0$, se cumplirá que $\varphi(u(t, \epsilon)) = u(t, \epsilon) > 0$ para $t \in (-\epsilon, 0)$. Por lo tanto, $(x(t, \epsilon), u(t, \epsilon)) \notin S$ si $\epsilon \neq 0$, implicando así que $(y, v) \notin C_S(x_0, u_0)$.

Conclusiones

En este trabajo probamos que, para el problema isoperimétrico de Lagrange con puntos extremos fijos y sujeto a restricciones de igualdad y desigualdad, la relación entre los distintos tipos de normalidad que se pueden encontrar en la literatura, algunas de sus caracterizaciones, y los multiplicadores de Lagrange asociados a una solución local, es similar a la que existe en los problemas de optimización en espacios de dimensión finita.

Recordemos que en estos problemas se busca minimizar un funcional de la forma $I(z) = \int_T L(t, z)dt$ (donde $z = (x, \dot{x})$ en cálculo de variaciones y $z = (x, u)$ en control óptimo) sobre un conjunto S compuesto por los elementos (x, u) que satisfacen una dinámica dada por una ecuación diferencial ordinaria de la forma $\dot{x}(t) = f(t, x(t), u(t))$, la condición de los puntos extremos fijos y restricciones isoperimétricas de igualdad y desigualdad.

Si una solución z_0 es normal con respecto al conjunto S, se dice que es débilmente normal, condición equivalente a una del tipo Mangasarian-Fromovitz que garantiza la existencia de multiplicadores de Lagrange asociados a z_0, pero no la unicidad. Si se conoce un multiplicador de Lagrange μ, correspondiente a z_0, este multiplicador es único si y sólo si z_0 es normal con respecto a S_1, subconjunto de S determinado por el signo de las componentes de μ; esta normalidad es también equivalente a una condición del tipo Mangasarian-Fromovitz estricto, caracterización que permitió en [9] probar que la regularidad es una consecuencia de la normalidad, obteniendo así condiciones de segundo orden en el cono $R_{S_1}(x_0)$.

Sin embargo, para que estas conclusiones fueran válidas aun considerando otro funcional, es decir, variando la función L en $\mathcal{F}(z_0)$, z_0 debe satisfacer la condición de normalidad fuerte (normalidad con respecto a S_0) la cual es equivalente a que las primeras variaciones de las restricciones de igualdad y de índices activos en x_0 sean linealmente independientes en el conjunto de direcciones admisibles.

En cambio, si el problema está definido por restricciones puntuales, muchas de las implicaciones anteriores ya no son válidas. La normalidad con respecto a S ($S_1(\mu)$), en general no se puede caracterizar por condiciones del tipo Mangasarian-Fromovitz (estricto) debido a que el comportamiento de las funciones conjunto valuadas $I_a(u_0(t))$, $\Gamma_+(t)$ y $\Gamma_0(t)$ puede variar dependiendo de los valores de t en el intervalo de tiempo. Si se conoce un multiplicador de Lagrange (p, μ), la normalidad con respecto a $S_1(\mu)$ garantiza la unicidad de dicho multiplicador, pero ya no es una equivalencia. En este contexto, hemos probado que la unicidad está caracterizada por una condición intermedia entre la normalidad débil y la normalidad con respecto a S_1. Por otro lado, la normalidad fuerte sí es una caracterización de la unicidad del multiplicador de Lagrange asociado a z_0 para $F \in \mathcal{F}(z_0)$ arbitraria.

Es fácil ver que los resultados obtenidos pueden ser extendidos a problemas más generales, como el de Bolza (ver [31]), donde el punto final $x(t_1)$ es variable, pues las condiciones de transversalidad que aparecen en el Principio Máximo de Pontryagin (Teorema 3.1.1) no interfieren en las demostraciones de dichos resultados.

Los métodos empleados en [5] para probar condiciones de segundo orden no se pueden aplicar a esta clase de problemas. De hecho, mostramos por medio de un ejemplo que la normalidad de z_0, aun si es del tipo fuerte, no es suficiente para garantizar que $\mathcal{R}_S(z_0) = \mathcal{T}_S(z_0)$ o que $\mathcal{R}_S(z_0) = C_S(z_0)$. Esto se debe a que, en los casos anteriores, los índices inactivos pueden ser completamente ignorados pues en una vecindad de z_0 éstos permanecen inactivos y no es necesario considerarlos; en cambio, como se ilustró en el Ejemplo 3.5.3, aquí representaron una parte fundamental del problema pues, dado que no se imponen condiciones para los elementos de $\mathcal{R}_S(z_0)$ en los intervalos de inactividad, fue posible encontrar una dirección $(y, v) \in \mathcal{R}_S(z_0)$ para la que no hay sucesiones de procesos en S convergiendo a z_0 en la dirección de (y, v). Sin embargo, aunque se encontraron ejemplos de soluciones fuertemente normales tales que $J(z_0, p, \mu; y, v) < 0$ para algún $(y, v) \in \mathcal{R}_S(z_0)$ y soluciones débilmente normales con $J(z_0, p, \mu; y, v) < 0$ para algún $(y, v) \in \mathcal{R}_{S_0}(z_0)$, ninguno de ellos contradice la conjetura sobre condiciones de segundo orden en $\mathcal{R}_{S_1}(z_0)$, por lo que la pregunta sobre cuál es la relación entre la normalidad con respecto a S_1 y la unicidad del multiplicador de Lagrange (p, μ) con el signo de $J(z_0, p, \mu; \cdot, \cdot)$ en $\mathcal{R}_{S_1}(z_0)$ continúa abierta.

Bibliografía

[1] Aubin JP, Frankowska H (1990) *Set-Valued Analysis*, Birkhäuser, Boston.

[2] Arutyunov AV, Pereira FL (2006) *Second-order necessary optimality conditions for problems without a priori normality assumptions*, Mathematics of Operations Research, **31**: 1-12.

[3] Arutyunov AV, Vereshchagina YS (2002) *On necessary second-order conditions in optimal control problems*, Differential Equations, **38**: 1531-1540.

[4] Bazaraa MS, Sherali HD, Shetty CM (1993) *Nonlinear Programming - Theory and Algorithms*, John Wiley, New York.

[5] Becerril JA (2018) *Normalidad Débil para Condiciones de Segundo Orden en el Problema Isoperimétrico de Lagrange*, Tesis Doctoral, UNAM.

[6] Becerril JA, Cortez KL, Rosenblueth JF (por aparecer) *Critical cones for regular controls with inequality constraints*. International Journal of Mathematical Analysis.

[7] Becerril JA, Cortez KL, Rosenblueth JF *Uniqueness of multipliers in optimal control: the missing piece*, IMA Journal of Mathematical Control & Information, dny033, https://doi.org/10.1093/imamci/dny033

[8] Becerril JA, Rosenblueth JF (2017) *Necessity for isoperimetric inequality constraints*, Discrete and Continuous Dynamical Systems Series A, **37**: 1129-1158.

[9] Becerril JA, Rosenblueth JF (2017) *The importance of being normal, regular and proper in the calculus of variations*, Journal of Optimization Theory & Applications, **172**: 759-773.

[10] Ben-Tal A (1980) *Second-order and related extremality conditions in nonlinear programming*, Journal of Optimization Theory & Applications, **31**: 143-165.

[11] Bigi G, Castellani M (2004) *Uniqueness of KKT multipliers in multiobjective optimization*, Applied Mathematics Letters, **17**: 1285-1290.

[12] Bonnans JF (1994) *Local analysis of Newton-type methods for variational inequalities and nonlinear programming*, Applied Mathematics & Optimization, **29**: 161-186.

[13] Bonnans JF, Osmolovskiĭ NP (2010) *Quadratic growth conditions in optimal control problems*, Contemporary Mathematics, **514**: 85-98.

[14] Clarke F (2013) *Functional Analysis, Calculus of Variations and Optimal Control*, Springer-Verlag, London.

[15] Cortez KL, Rosenblueth JF (2016) *A second order constraint qualification for certain classes of optimal control problems*, WSEAS Transactions on Systems and Control, **11**: 419-424.

[16] Cortez del Río KL, Rosenblueth JF (2017) *Extended critical directions for time-control constrained problems*, International Journal of Circuits, Systems and Signal Processing, **11**:1-11.

[17] Cortez KL, Rosenblueth JF (2018) *Normality and uniqueness of Lagrange multipliers*, Discrete & Continuous Dynamical Systems Series A, **38**: 3169-3188.

[18] Courant R (1962) *Calculus of Variations*, Courant Institute of Mathematical Sciences, New York University, NY.

[19] de Pinho MR, Rosenblueth JF (2007) *Mixed constraints in optimal control: an implicit function theorem approach*, IMA Journal of Mathematical Control & Information, **24**: 197-218.

[20] Deng S (1996) *On uniqueness of Lagrange multipliers in composite optimization*, Journal of Mathematical Analysis & Applications, **201**: 689-696.

[21] Fujiwara O, Han SP, Mangasarian OL (1984) *Local duality of nonlinear programs*, SIAM Journal of Control & Optimization, **22**: 162-169.

[22] Gilbert EG, Bernstein DS (1983) *Second-order necessary conditions in optimal control: accesory-problem results without normality conditions*, Journal of Optimization Theory & Applications, **41**: 75-106.

[23] Giorgi G, Guerraggio A, Thierfelder J (2004) *Mathematics of Optimization: Smooth and Nonsmooth Case*, Elsevier, Amsterdam.

[24] Hestenes MR (1966) *Calculus of Variations and Optimal Control Theory*, John Wiley, New York.

[25] Hestenes MR (1975) *Optimization Theory: The Finite Dimensional Case*, John Wiley, New York.

[26] Kyparisis J(1985) *On uniqueness of Kuhn-Tucker multipliers in nonlinear programming*, Mathematical Programming, **32**: 242-246.

[27] Kyparisis J (1990) *Sensitivity analysis for nonlinear programs and variational inequalities with non-unique multipliers*, Mathematics of Operations Research, **15**: 286-298.

[28] Lempio F, Maurer H (1980) *Differential stability in infinite-dimensional nonlinear programming*, Applied Mathematics & Optimization, **20**: 139-152.

[29] Levitin E, Milyutin A, Osmolovskii NP (1978) *Conditions of high order for a local minimum for problems with constraints*, Russian Math Surveys, **33**: 97-178.

[30] Loewen PD, Zheng H (1994) *Generalized conjugate points in optimal control*, Proceedings of the 33rd IEEE Conference on Decision & Control, Lake Buena Vista, Florida, **4**: 4004-4008.

[31] Loewen PD, Zheng H (1994) *Generalized conjugate points for optimal control problems*, Nonlinear Analysis, Theory, Methods & Applications, **22**: 771-791.

[32] Malanowski K (2003) *On normality of Lagrange multipliers for state constrained optimal control problems*, Optimization, **52**: 75-91.

[33] Maurer H, Oberle HJ (2002) *Second order sufficient conditions for optimal control problems with free final time: the Riccati approach*, SIAM Journal on Control & Optimization, **41**: 380-403.

[34] McShane EJ (1973) *The Lagrange multiplier rule*, The American Mathematical Monthly, **80**: 922-925.

[35] Milyutin AA, Osmolovskii NP (1998) *Calculus of Variations and Optimal Control*, Translations of Mathematical Monographs, 180, American Mathematical Society, Providence, Rhode Island.

[36] OSMOLOVSKII NP (1975) *Second order conditions for a weak local minimum in an optimal control problem (necessity, sufficiency)*, Soviet Math Dolk, **16**: 1480-1484.

[37] ROSENBLUETH JF (2007) *Convex cones and conjugacy for inequality control constraints*, Journal of Convex Analysis, **14**: 361-393.

[38] ROSENBLUETH JF (2015) *Modified critical directions for inequality control constraints*, WSEAS Transactions on Systems & Control, **10**: 215-227.

[39] ROSENBLUETH JF, SÁNCHEZ LICEA G (2013) *Cones of critical directions in optimal control*, International Journal of Applied Mathematics and Informatics, **7**: 55-67.

[40] RUSSAK IB (1975) *Second order necessary condition for problems with state inequality constraints*, SIAM Journal on Control, **13**: 372-388.

[41] RUSSAK IB (1975) *Second order necessary condition for general problems with state inequality constraints*, Journal of Optimization Theory & Applications, **17**: 43-92.

[42] SHAPIRO A (1997) *On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints*, SIAM Journal of Optimization, **7**: 508-518.

[43] STEFANI G, ZEZZA PL (1996) *Optimality conditions for a constrained control problem*, SIAM Journal on Control & Optimization, **34**: 635-659.

[44] WACHSMUTH G (2013) *On LICQ and the uniqueness of Lagrange multipliers*, Operations Research Letters, **41**: 78-80.

[45] WARGA J (1996) *A second-order Lagrangian condition for restricted control problems*, Journal of Optimization Theory & Applications, **24**: 475-483.

[46] ZEIDAN V (1996) *Admissible directions and generalized coupled points for optimal control problems*, Nonlinear Analysis, Theory, Methods & Applications, **26**: 479-507.